DOCUMENT RESUME

ED 199 295 TM 810 277

AUTHOR	McCaffrey, Arthur
TITLE	Testing a Model of Communicative Competence in the Classroom. Final Report.
INSTITUTION	Judge Baker Guidance Center, Boston, Mass.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	80
GRANT	NIE-G-76-0042
NOTE	297p.
EDRS PRICE	MF01/PC12 Plus Postage.
DESCRIPTORS	Classroom Communication; Cognitive Processes; *Communicative Competence (Languages); Diagnostic Teaching; Diagnostic Tests; *Evaluation Methods; Grade 2; *Intervention; Learning Activities; *Models; Primary Education; *Program Effectiveness; *Psycholinguistics; Underachievement
IDENTIFIERS	*Peer Modeling

ABSTRACT

The results of a study in applied educational psycholinguistics are presented. The project used a theoretical model of communicative competence to develop testing techniques for diagnosing levels of ability in functional language skills, and to develop training techniques to promote the acquisition of speaking and listening skills among elementary school children in three urban metropolitan schools. Three main issues of the study were: (1) Evaluation: a theoretical analysis and description of the communication process was made, and a battery of perceptual, cognitive and linguistic measures were constructed which reliably and validly trichotomised a sample of grade 2 students into low, moderate, and high levels of oral communication competence; (2) Intervention/Training: the evaluation data were used to create an experimental group of low competent subjects who were enrolled in a communication training program to improve their expressive language skills; and (3) Assessment of Training: statistical comparisons were made of experimental versus control subjects and a number of significant training and modeling effects were found on a variety of cognitive and pragmatic-linguistic measures. Two major language functions were discovered--spatio-temporal language, and logical language--and their implications for cognitive and linguistic development were examined. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED199295

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM | CS
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

FINAL REPORT

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

NATIONAL INSTITUTE OF EDUCATION
PROJECT NO. G-76-0042

TESTING A MODEL OF COMMUNICATIVE COMPETENCE IN THE CLASSROOM

Arthur McCaffrey, Ph.D.

Judge Baker Guidance Center
Children's Hospital
295 Longwood Avenue
Boston, Massachusetts 02115

1980

TM 810 277

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

ERIC Summary

A. McCaffrey. Testing a Model of Communicative Competence in the Classroom. Final Report, National Institute of Education Project No. G-76-0042. Judge Baker Guidance Center, Children's Hospital, Boston, Mass., 1980.

The report presents the results of a study in Applied Educational Psycholinguistics. The project used a theoretical model of communicative competence to (1) develop testing techniques for diagnosing levels of ability in functional language skills, and (2) develop training techniques to promote the acquisition of speaking and listening skills among elementary schoolchildren in 3 urban metropolitan schools.

The three main issues of the study were: I. Evaluation: a theoretical analysis and description of the communication process was made, and a battery of perceptual, cognitive and linguistic measures were constructed which reliably and validly trichotomised a sample of second-graders into Low, Moderate, and High levels of oral communication competence; II. Intervention/Training: the Evaluation data were used to create an experimental group of Low competent Ss. who were enrolled in a communication training program to improve their expressive language skills. The program was based on a curriculum of communication games. These were derived from a set of psycho-pedagogical principles which are presented to show teachers how to create language-learning environments in the classroom which will afford children practice in functional language usage, and the beneficial role of peer models in such exercises is examined; and III. Assessment of Training: statistical comparisons were made of experimental vs. control Ss. and and a number of significant Training and Modeling effects were found on a variety of cognitive and pragmatic-linguistic measures. Two major language functions were discovered-- Spatio-temporal Language, and Logical Language-- and their implications for cognitive and linguistic development were examined. The developmental, educational and research implications of the study's findings are discussed.

ABSTRACT

Testing a Model of Communicative Competence in the Classroom

This report presents the results of a study in Applied Educational Psycholinguistics. The aim of the project was to study the effectiveness of a theoretical model of the communication process for (1) generating testing techniques which would diagnose oral communication abilities, and (2) developing training techniques which would foster the acquisition of speaking and listening skills among elementary schoolchildren.

The report is in 3 parts dealing with the main issues of the study:
A. Evaluation: the theoretical analysis and practical measurement of communicative competence in school-age children.
B. Intervention/Training: use of the theoretical rationale and Evaluation data to create a communication curriculum and learning environment for training and developing pragmatic-linguistic skills in the classroom.
C. Assessment of Training Effects: empirical assessment of the effectiveness of the Training Program in fostering the development of oral communication abilities of less competent students.

A brief outline of the three sections of the study is as follows:

A. Evaluation

Before one intervenes, one has to measure, and before one measures, one has to describe. Part I of this report deals with the issues of description and measurement of communicative competence. The position is taken that the act of constructing a comprehensible oral communication for another person is not just a linguistic phenomenon, but is also a social and cognitive event. A theoretical model is presented which proposes three basic components of communicative competence — perceptual, cognitive, and linguistic — and the psychological and developmental aspects of these processes are discussed with reference to speaker-listener dialogue.

The practical, functional descriptions of communication behavior provided by the model are related to recent developments in pragmatics research in developmental psycholinguistics. The model also provides a frame of reference for a review and critique of studies of communication in children, and educational implications are considered.

The theoretical model of communicative competence provides a rationale for both the evaluation and the training of speaking and listening competences in schoolchildren. For the experimental application of the model in the first, Evaluation phase of the study, a battery of perceptual, cognitive and verbal measures was constructed and administered to approximately 150 second-grade pupils in public schools in a metropolitan area. An elaborate scoring system was designed to analyse the response data for the three component areas of communication. For the verbal-communication data in particular, an elaborate Propositional Analysis Scheme was developed to analyse important structural and qualitative factors in oral message construction.

All test data were scored, coded, keypunched and submitted to computer analyses. The results were used to determine the competency status of each subject, and the total sample was trichotomised into three levels of ability -- Low, Moderate, and High competence. Statistical analyses were made to establish significant differences between these three groups on the major variables of analysis from the three domains of perceptual, cognitive and linguistic functioning. In addition, factor analyses of the verbal message data further sub-divided the language domain into two functional categories of language use — Spatio-Temporal Language (STL) and Logical Language (LL).

The final section of the Evaluation phase of the study concentrated on the group of Low Competent Ss, in preparation for the next phase of the research, Training of communicative competence. These Ss were randomly assigned to experimental and control groups and four such groups were formed — three Training and one Control. The three Training groups were further differentiated by level of peer-modeling within each Training group. The target of the training sessions were the Low Ss in each group, but in group I these Ss were partnered with other Low competent peers; in group II they were partnered with Moderate competent Ss; and in group III the Low Ss were paired with High competent peer models. Statistical analyses were performed to establish the initial equivalence of all four experimental groups on the major measures prior to training.

B. Intervention

Part II of the report deals with the application of the model and its test results to the analyses of pragmatic language learning in the classroom. Following the diagnoses of communication ability made in Part I, the second phase of the study establishes a philosophy, pedagogy, and curriculum for fostering the expressive language abilities of the Low competent trainees. The three major topics addressed in this classroom application of the study are a) the creation and structuring of appropriate language environments in the classroom; b) the role of training and practice for developing a well-rounded competence across a variety of situations, purposes and language uses; c) the role of more competent peer models in facilitating the learning of new skills by the trainees.

The oral communication program was developed and implemented over a six-month period of the school year, and the day-to-day workings of the curriculum are described. A set of psycho -pedagogical principles are presented for use in any program where the teacher's goal is the development of functional language competencies. The role of these principles in the training program is discussed and examples of communication games and exercises are provided. Part II concludes with a series of recommendations for future uses of the program in the classroom by teachers.

C. Assessment of Training Effects

Part III of the report presents the empirical evidence for the effectiveness of the program in developing the competence of the experimental trainees, as measured by the battery of perceptual, cognitive and linguistic measures.

Exploratory analyses were made of the post-training data for the experimental and control Ss, in order to identify important trends and patterns in the data. On the basis of these, a further set of confirmatory analyses were made using correlational, factor analytic and multivariate techniques. The results showed a variety of significant Training and Pairing effects for various measures in the Cognitive, STL, and LL domains. Interrelationships between the domains of competence are discussed and suggestions are made for modifying the theoretical model in the light of the empirical evidence. Based on the successes of the program, an appraisal is made of the utility of the model (and its accompanying measurement and training techniques) for the diagnosis and remediation of underdeveloped competencies by teachers, clinicians, researchers and other professional practitioners.

The study data also indicated important relations between pragmatic-linguistic and cognitive functions, and the findings spotlight the mediating roles played both by cognitive processes and the peer-models in transmitting learning effects from the communication training program. For future research, special emphasis is placed on the importance of language measures which deal with both the form and the content of children's verbal messages and which are sensitive to the part-whole, subordinate-superordinate relations involved in the construction of such messages.

TABLE OF CONTENTS

ERIC Summary i
Study ABSTRACT ii
List of TABLES viii
List of FIGURES x
List of APPENDICES x

1. INTRODUCTION 1

A. Evaluation
B. Intervention
C. Assessment of Training

2. Background and General Purpose of the Study 3

3. Educational Outcomes 7

PART I: EVALUATION OF COMMUNICATIVE COMPETENCE

I. THEORY 10

1. Definitions and Limitations 10
2. A Model of Communicative Competence 11
3. A Critique of Selected Studies 15
4. Some experimental tests of the Model 16

II. PRACTICE

1. Sample 18
2. Test Battery 19
  i. Perceptual Tests 19
  ii. Cognitive Tests 20
  iii. Language Tests 23
    A. Toy Village/Action-Sequence Tests 24
    B. Picture Description Tests 26
3. Data Management 29
4. Scoring Systems and Major Variables of Analysis 29
  i. Perceptual Measures 29
  ii. Cognitive Measures 30
  iii. Language Measures: A Propositional Analysis Scheme 30
  iv. Language Measures: The Major Variables 34
    a) Action-Sequence 35
    b) Picture Description 37
  v. Reliability of Scoring 39

III. RESULTS 41

1. Categorisation of Competency Status (COMPSTAT) 41
2. Validating the Trichotomy of Competence 42
    i. Univariate Analyses 42
    ii. Discriminant Function Analysis 47
3. Interrelationships Within and Between Domains 51
    i. Language: Item Analysis and Canonical Correlation 51
    ii. Language: Quality/Quantity relations 52
    iii. Perception: Correlations 54
    iv. Cognition: Correlations 54
    v. Inter-Domain relations 54
4. Identifying Two Major Language Functions 56
    i. Factor Analyses 56
5. External Validation of the COMPSTAT Scale 58
6. Summary Discussion of the Evaluation Results 59

IV. FROM EVALUATION TO INTERVENTION 62

1. Selecting the Experimental Groups (EXPGROUPS) and Peer Models 62
2. Establishing the Initial Equivalence of the EXPGROUPS 64
3. The Experimental Predictions 64

PART II: TRAINING OF COMMUNICATIVE COMPETENCE

1. Overview of the Training Program 66
    i. Major characteristics of the Training Model 66
2. The Structure of the Training Program 71
    i. Formation of the Experimental Groups 71
    ii. Logistics of the Training Program 72
3. Philosophy and Content of the Training Program 78
    i. Some opinions and assumptions 78
    ii. The Basic Principles 80
4. Discussion 84
    i. The effectiveness of the Program 84
        a) Philosophy, Content and Structure of the Training Model 85
        b) The Psycholinguistic Content of an Oral Language Curriculum 87
        c) Learning effects and generalisation 88
        d) The general applicability of the Program and the issue of Classroom Management 89
        e) Monitoring and Evaluating a Training Program — short vs. long-term goals 91
5. Summary 93

PART III: ASSESSMENT OF TRAINING EFFECTS

I. POST-TRAINING EVALUATIONS 94

1. Testing Procedures 94
2. Data Management 95

II. RESULTS FOR PRIMARY MEASURES: EXPLORATORY ANALYSES 98

1. Perceptual Domain 98
2. Cognitive Domain 102
3. Language Domains 108
   A. Spatio-temporal Language (STL) 108
   B. Logical Language (LL) 120
   C. Summary Discussion and Implications of the Language Domain Results 131

III. RESULTS FOR PRIMARY MEASURES: CONFIRMATORY ANALYSES 136

1. Post-training Intercorrelations 136
2. Factor Analyses 141
3. Confirmatory Multivariate Analyses 143
   i. Introduction 143
   ii. Overview of the Analyses 144
   iii. Results 147
      a) A General Test of the Model 147
      b) Perception and Cognition 148
      c) Cognitive Gains 149
      d) The STL Domain 150
      e) The LL Domain 151
      f) Cognition and Language 152
      g) The Two Language Domains and the Quality/Quantity Issue 154
      h) Grand Manova 156
      i) Summary Discussion of Manova Results 157
4. Summary and Conclusions 159
5. Recommendations 160

LIST OF TABLES

1-a. Compstat Composition of Test Dyads for both Language Domains, PRETEST.
1-b. " " " " " " , POSTTEST.
1-c. Breakdown of Test Dyad Pairings x EXPGROUP x COMPSTAT x STL, POSTTEST.
1-d. " " " " " " x LL, POSTTEST.
2. PERCEPTION: Crossbreaks and One-way Anova for ACCURACY x EXPGROUPS, PRE/POST.
3. " " " " " SPEED x EXPGROUPS, PRE/POST.
4. " " " " " SPEED x ACCURACY x EXPGROUPS, PRE/POST.
5. " " " " " SPEED x INACCURACY x EXPGROUPS, PRE/POST.
6. PERCEPTION: Crossbreaks and One-way Anova for COMPOSITE ACCURACY x EXPGROUPS, PRE/POST.
7. PERCEPTION: Intercorrelations among Measures, PRE/POST.
8. COGNITION: Crossbreaks and One-way Anova for HOUSE CONSTRUCTION TASK x EXPGROUPS, PRE/POST.
9. " " " " " HOUSE PHOTOS TASK x EXPGROUPS, PRE/POST.
10. " " " " " MOUNTAIN PERSPECTIVES TASK x EXPGROUPS, PRE/POST.
11. " " " " " 1-SCORES x EXPGROUPS, PRE/POST.
12. " " " " " 2-SCORES x EXPGROUPS, PRE/POST.
13. " " " " " 3-SCORES x EXPGROUPS, PRE/POST.
14. " " " " " SUMMARY MEASURE (COGSUM) x EXPGROUPS, PRE/POST.
15. " Intercorrelations among Measures, PRE/POST.
16. LANGUAGE: STL, Summary Raw Data for EXPGROUPS, PRE/POST.
17. " STL, Crossbreaks and Two-way Anova for PREP. USAGE x EXPGROUPS, PRE/POST.
18. " " " " " " QUANTITY x EXPGROUPS, PRE/POST.
19. " " " " " " QUALITY x EXPGROUPS, PRE/POST.
20. " " " " " " EFFICIENCY I x EXPGROUPS,PRE/POST.
21. " " , Crossbreaks and One-way Anova for EFFICIENCY II GAIN SCORES, for EXPGROUPS, POSTTEST.
22. " STL, Crossbreaks and Two-way Anova for EFFICIENCY III x EXPGROUPS, PRE/POST.
23. " " " " " " AMBIGUITY x EXPGROUPS, PRE/POST.
24. " " " " " " INEFFICIENCY I x EXPGROUPS, PRE/POST.
25. " " " " " " INEFFICIENCY II x EXPGROUPS, PRE/POST.
26. " " " " " " ORGANISATION x EXPGROUPS, PRE/POST.
27. " " " " " " COMPOSITE COMPETENCE x EXPGROUPS, PRE/POST.
28. " " " " " " COMPOSITE EFFICIENCY II x EXPGROUPS, PRE/POST.
29. " STL, Crossbreaks and One-way Anova for COMPOSITE COMPETENCE GAIN SCORES x EXPGROUPS, POSTTEST.
30. " STL Crossbreaks and Two-way Anova for PREDICTED QUALITY x EXPGROUPS, PRE/POST.

Tables (cont.)/2

31. LANGUAGE: LL, Summary Raw Data by Tasks, for EXPGROUPS, PRE/POST.
32. " LL, Crossbreaks and Two-way Anova for QUANTITY x EXPGROUPS, PRE/POST.
33. " " " " " " DISCRIMINABILITY INDEX x EXPGROUPS, PRE/POST.
34. " " " " " " QUALITY x EXPGROUPS, PRE/POST.
35. " " " " " " EFFICIENCY x EXPGROUPS, PRE/POST.
36. " " " " " " INEFFICIENCY x EXPGROUPS, PRE/POST.
37. " " " " " " ORGANISATION x EXPGROUPS, PRE/POST.
38. " " " " " " PREDICTED QUALITY x EXPGROUPS, PRE/POST.
39. PERCEPTION: Crossbreaks and One-way Anovas for ACCURACY AND SPEED x COMPSTAT, PRETEST.
40. COGNITION: Crossbreaks and One-way Anovas for TASKS, SCORES, and SUMMARY MEASURE (COGSUM) x COMPSTAT, PRETEST.
41. LANGUAGE: STL, Summary Raw Data, by COMPSTAT, PRETEST.
42. " : STL, Crossbreaks and Two-way Anovas for COMPSTAT x TEST POSITION, PRETEST.
43. LANGUAGE: LL, Summary Raw Data, by COMPSTAT, PRETEST.
44. " LL, Crossbreaks and Two-way Anovas for COMPSTAT x TEST POSITION, PRETEST.
45. DISCRIMINANT FUNCTION CLASSIFICATION ANALYSIS, for COMPSTAT, PRETEST.
46. REGRESSION ANALYSIS: Correlation of QUALITY and QUANTITY for two Language Domains, PRETEST.
47. Inter-Domain Correlations, PRETEST.
48. " " " , POSTTEST.
49. FACTOR ANALYSES of Language Domain Variables, PRETEST.
50. " " " " " , POSTTEST.

M1. MANOVAS of Training Effects for Four Variables from each Domain, for EXPGROUPS.
M2. " " " " " for GAIN SCORES x EXPGROUPS.
M3. " " " for PERCEPTION & COGNITION, for EXPGROUPS.
M4. " " " " " " , with the LL DISCRIMINABILITY Variable, for EXPGROUPS.
M5. MANOVAS of Training Effects for 5 STL Variables, for EXPGROUPS.
M6. " " " " 4 LL " " " .
M7. " " " " COGNITIVE & STL Variables, for EXPGROUPS.
M8. " " " " " " LL " " " .
M9. " " " " for STL & LL EFFICIENCY Measures x EXPGROUPS.
M10. " " " " STL & LL QUALITY/QUANTITY Measures x EXPGROUPS.
M11. " " " " " " " PREDICTED QUALITY Measures x EXPGROUPS.
M12. " " " , GRAND MANOVA for all DOMAINS, for EXPGROUPS.

LIST OF FIGURES

1. Cognitive Task #1 --- House Construction.

2. Cognitive Task #2 --- House Pictures.

3. Cognitive Task #3 --- Mountains.

4. Experimental Design for Training: EXPGROUPS x COMPSTAT. (p. 62)

5. Peer Modeling and Practice in the 4 Treatment Groups. (p. 63)

LIST OF APPENDICES

A. Perceptual Tasks Score Sheet.
B. Cognitive Tasks Score Sheet.
C. Language Tasks: STL, Comprehension of Prepositions Test.
D. " " : STL Test Items.
E. " " : STL Answer Sheet.
F. " " : LL Test Items/Answer Sheet.
G. " " : Excerpts from Test Transcripts, showing responses to STL and LL Tasks.

II A. A Core Curriculum of Communication Games.

TESTING A MODEL OF COMMUNICATIVE COMPETENCE IN THE CLASSROOM

ARTHUR Mc CAFFREY

TESTING A MODEL OF COMMUNICATIVE COMPETENCE IN THE CLASSROOM

1. Introduction

This Report presents the results of a research project in applied psycholinguistics, dealing with an experimental intervention into the language learning of schoolchildren. The major aim of the project was to study the effectiveness of a psycholinguistic model of the communication process for generating testing and training techniques which would be used both to diagnose and foster the oral communication skills of elementary schoolchildren.

The main issues in the study are Evaluation, Intervention, and Assessment of Training effects, and these are the major divisions of a research program which combines linguistic and educational concerns. In the rationale that will be developed below a case will be made for studying the acquisition of language _usage_ skills during the early school years rather than the acquisition of linguistic _knowledge_ per se. The study is motivated by an awareness of the marked discrepancy that often exists between what the pre-school child has learned about the structure of his language and what he has learned about the ways of using his linguistic repertoire for effective oral communication. Despite the very worthy and very democratic Chomskyan thesis that all children, regardless of race, creed, class, or color, are born with the capacity to acquire language, it is still a fact that children arrive at school showing very wide, very obvious, and non-egalitarian differences in their ability to use whatever it is they have acquired about language for the practical purposes of interpersonal communication.

These differences in communicative competence (or linguistic 'savoir faire') between children immediately raise important developmental and educational questions concerning a) the _nature_ of these individual differences in language usage skills, and b) the possible _antecedents_ of these differences. The _evaluation phase_ of this study (Part I below) addresses itself to the first problem of describing and assessing these differences in competence. The second question is tackled through the _training phase_ of the study (Part II below) which attempts to counteract any adverse pre-school language experiences by manipulating the language environment of the classroom via educational intervention strategies. Our ultimate pedagogical concern, therefore, is with the problem of how school-age children can be taught language usage skills by techniques that a) build upon the pre-acquired language repertoire the child brings to school with him and b) enhance the expansion of this repertoire by providing him with learning experiences that require him to utilize his linguistic knowledge for practical, purposeful communication.

The three sections of the study will be outlined briefly.

A. Evaluation

Before one intervenes, one has to measure, and before one measures, one has to describe. Part I of this report deals with the issues of description and measurement of communicative competence. The position is taken that the act of constructing a comprehensible oral communication for another person is not just a linguistic phenomenon, but is also a social and cognitive event. A theoretical model is presented which proposes three basic components of communicative competence -- perceptual, cognitive, and linguistic -- and the psychological and developmental aspects of these processes are discussed with reference to speaker-listener dialogue.

The practical, functional descriptions of communication behavior provided by the model are related to recent developments in pragmatics research in developmental psycholinguistics. The model also provides a frame of reference for a review and critique of studies of communication in children, and educational implications are considered.

The theoretical model of communicative competence provides a rationale for both the evaluation and the training of speaking and listening competences in schoolchildren. For the experimental application of the model in the first, Evaluation phase of the study, a battery of perceptual, cognitive and verbal measures was constructed and administered to approximately 150 second-grade pupils in public schools in a metropolitan area. An elaborate scoring system was designed to analyse the response data for the three component areas of communication. For the verbal-communication data in particular, an elaborate Propositional Analysis Scheme was developed to analyse important structural and qualitative factors in oral message construction.

All test data were scored, coded, keypunched and submitted to computer analyses. The results were used to determine the competency status of each subject, and the total sample was trichotomised into three levels of ability -- Low, Moderate, and High competence. Statistical analyses were made to establish significant differences between these three groups on the major variables of analysis from the three domains of perceptual, cognitive and linguistic functioning. In addition, factor analyses of the verbal message data further sub-divided the language domain into two functional categories of language use -- Spatio-Temporal Language (STL) and Logical Language (LL).

The final section of the Evaluation pahse of the study concentrated on the group of Low Competent Ss, in preparation for the next phase of the research, Training of communicative competence. These Ss were randomly assigned to experimental and control groups and four such groups were formed -- three Training and one Control. The three Training groups were further differentiated by level of peer-modeling within each Training group. The target of the training sessions were the Low Ss in each group, but in group I these Ss were partnered with other Low competent peers; in group II they were partnered with Moderate competent Ss; and in group III the Low Ss were paired with High competent peer models. Statistical analyses were performed to establish the initial equivalence of all four experimental groups on the major measures prior to training.

Part I of the report concludes with a discussion of the experimental outcomes and predictions for the Training phase of the study.

B. Intervention

Part II of the report deals with the application of the model and its test results to the analyses of pragmatic language learning in the classroom. Following the diagnoses of communication ability made in Part I, the second phase of the study establishes a philosophy, pedagogy, and curriculum for fostering the expressive language abilities of the Low competent trainees. The three major topics addressed in this classroom application of the study are a) the creation and structuring of appropriate language environments in the classroom; b) the role of training and practice for developing a well-rounded competence across a variety of situations, purposes and language uses; c) the role of more competent peer models in facilitating the learning of new skills by the trainees.

The oral communication program was developed and implemented over a six-month period of the school year, and the day-to-day workings of the curriculum are described. A set of psycho -pedagogical principles are presented for use in any program where the teacher's goal is the development of functional language competencies. The role of these principles in the training program is discussed and examples of communication games and exercises are provided. Part II concludes with a series of recommendations for future uses of the program in the classroom by teachers.

C. Assessment of Training Effects

Part III of the report presents the empirical evidence for the effectiveness of the program in developing the competence of the experimental trainees, as measured by the battery of perceptual, cognitive and linguistic measures.

Exploratory analyses were made of the post-training data for the experimental and control Ss, in order to identify important trends and patterns in the data. On the basis of these, a further set of confirmatory analyses were made using correlational, factor analytic and multivariate techniques. The results showed a variety of significant Training and Pairing effects for various measures in the Cognitive, STL, and LL domains. Interrelationships between the domains of competence are discussed and suggestions are made for modifying the theoretical model in the light of the empirical evidence. Based on the successes of the program, an appraisal is made of the utility of the model (and its accompanying measurement and training techniques) for the diagnosis and remediation of underdeveloped competencies by teachers, clinicians, researchers and other professional practitioners.

The study data also indicated important relations between pragmatic-linguistic and cognitive functions, and the findings spotlight the mediating roles played both by cognitive processes and the peer-models in transmitting learning effects from the communication training program. For future research, special emphasis is placed on the importance of language measures which deal with both the form and the content of children's verbal messages and which are sensitive to the part-whole, subordinate-superordinate relations involved in the construction of such messages.

2. The Background and General Purpose of the Study

There has been much debate during the last two decades over the essential characteristics of good verbal communication skills, how these might be acquired by children, and how the modern grade school classroom might provide and foster the kinds of listening and speaking experiences that are the necessary prerequisities for the acquisition of communicative competence (Baldwin & Garvey, 1970; Baldwin, McFarlane, & Garvey, 1971; Bernstein, 1970; Cazden, 1970; Cazden, John, & Hymes, 1972; Flavell, 1968; Garvey, 1972; Heider, 1968; McCaffrey, 1970; Moffett, 1968; Bullock Report, 1975). For a long period of time, however, this debate over the nature and purpose of language education was carried on largely by educators and practitioners, and it did not find an echo in the basic research and theoretical developments that were taking place in psycholinguistics. Much research progress was made in identifying some of the processes underlying the acquisition of linguistic knowledge in the form of syntactic and semantic rules (Brown, 1973; Fillenbaum, 1971; McNeill, 1970), but much less attention was being paid to the child's acquisition of language usage skills. Consequently, translation of research findings into practical applications was frequently difficult, if not impossible.

Fortunately, recent changes in the emphases and fields of inquiry in psycholinguistic research have brought about a change of perspective that bodes well for the educator and practitioner. The research and theoretical focus has swung from syntagmatic-structural aspects of language development to pragmatic-functional concerns where situational and intentional constraints on language usage are the objects of study. Comprehensive reviews of the new pragmatics movement in psycholinguistics are available in the literature (e.g., Tamir, 1979; Gleason & Weintraub, 1978; McCaffrey, 1977). We will briefly review some of the more important developments in theory and research in this field in order to raise important questions about language education and the acquisition of communicative competence during the elementary school years.

While the new emphasis on the acquisition of language functions is welcome, one must ask how well the new theory and research in pragmatics helps us to conceptualise and operationalise fresh approaches to practical problems of language development and education. The recent concern with the nature and developmentof language functions in childhood is a hopeful sign for psychologists in the clinical and educational fields who have to deal with language at the levels of performance and use. Time alone will tell, however, if the new vocabulary of "speech acts" denotes more than just a superficial change in terminology and if it will provide the sought-for solutions to real problems in elementary language education.

At the moment, however, the signs are promising. Chomsky and "competence", in the psycholinguistics literature, are giving way to Austin (1962), Searle (1969), Grice (1967), and other philosophers of language who espouse "illocutionary acts", "communicative intentions", and the"propositional content" of messages in their choice of theoretical and research problems. This transition, therefore, from the earlier transformational period in psycholinguistics to the current pragmatic period indicates clearly that we appear to have moved, in a theoretical sense at least, from a level of concern with language "knowledge" (an epistemological concern that showed much commonality between Chomskyan and Piagetian philosophy) to a level of interest in how language is used in a social and interactional context (a functional concern that will hopefully also make for more commonality between the endeavors of the pragmatics theorists and those of applied psycholinguists).

This theoretical shift from "conceptions" to "actions" is illustrated by recent research on the nature and acquisition of "Speech Acts", on the relationships between grammatical forms and interactional goals, and on the role of prosodic and contextual features in determining message content and effect.

Dore (1975,1976), for example, has attempted to tackle the neglected, but thorny problem of the role of communicative intentions in determining the form and content of children's utterances. The problem for Dore is how 'speech acts' are acquired in infancy and early childhood, and how the child's pragmatic intentions (requesting, demanding, calling, protesting, negating, etc.) gradually become grammaticalised as as semantic and syntactic structures. Bruner (1975) also has been studying speech acts in early infancy as a way of bridging the gap between pre-linguistic perceptual and cognitive behaviors involved in attention and action, and the predicative and case grammatical forms of linguistic communication. Bruner suggests that "early language, to be acquired, must reflect the nature of the cognitive processes whose output it encodes." (1975, pp. 3-4). Some of the current research on the acquisition of language functions also provides a continuity with problems studied in previous syntactic research. Halliday (1973, 1975), for example, has theorised about the issue of "pragmatic universals", and has made a sociolinguistic analysis of seven basic and very common language functions acquired by children. This tentative system, he suggests, may be acquired in an ordered sequence over development and have universal characteristics. More recently, an ambitious but promising attempt has been made by Bates (1976) to integrate the study of speech acts and intentions with Piagetian cognitive-developmental theory. Bates has chosen to study such topics as the development of the child's use of propositions, the nature of his goal-directed performative utterances, the presuppositions that he learns to share with particular listeners, and his mastery of higher-order conversational postulates that will make him capable of more complex dialogue across a variety of contexts. Each of these pragmatic topics is explored by Bates in relation to each of the major pre-operational and operational stages of Piaget's system.

All of this new theory and research is very interesting and stimulating, but, from an applied-experimental point of view, there are some troublesome aspects surrounding this recent spurt of interest in pragmatics. First of all, a great deal of the new theorising and speculation about speech acts is based on data obtained with very small samples of subjects, and questions have to be raised about the generalisability of the results. Secondly, just as it did during the Chomskyian revolution, psycholinguistics has gone outside of the field of psychology to search for its pragmatics theory among philosophers of language. And just as uncritically, much of this borrowed theory has been imported wholesale with insufficient empirical scrutiny or testing. A third criticism is a developmental one: the age range covered in recent pragmatics research is overly restricted and reflects the same lop-sided preoccupation with the language acquisition of the infant and pre-school child that has charactised much of the developmental psycholinguistic research of the last ten years. What about the development of speech acts in the older, school-age child who is already at the stage of concrete operations? There can and should be a profitable interface between basic pragmatics research on the mastery of language functions and applied psycholinguistic research which would have direct implications for oral language curricula in elementary education.

Given this need, therefore, the focus of the present study will be on what Rommetveit (1968) has called the "pragmatic" aspects of language learning which have to do with the conditions and effects of expressive language use among schoolchildren. That is, while we are aware of the important contributions made by studies of the acquisition of grammatical structure, we choose to emphasise that other factors besides structural ones must be looked at in order to account for known differences in language performance between children entering school. Educational researchers must take more account of the interaction between structural and functional linguistic processes, and tackle the question of how the child uses what he has acquired. As Slobin (1968) remarks:

> Each child is equipped with a basic, strictly linguistic competence which can be differentially shaped to carry out a variety of sociolinguistic functions.

Of the many co-determining environmental influences involved in this "shaping" process and the child's acquisition of communicative competence, not the least of them is the child's socio-economic background and the kind of "models" of language usage he is exposed to in the home and in the neighborhood, through parents and peers (Bernstein, 1970; Heider, 1968). Structural processes in language acuisition may very well have innate components (Chomsky, 1968), but children have to learn the functions of language through a process of social interaction. One view of this kind of learning is offered by Hymes:

> The acquisition of competency for use, indeed, can be stated in the same terms as the acquisition of competence for grammar. Within the developmental matrix in which knowledge of the sentences of a language is acquired, children also acquire knowledge of a set of ways in which sentences are used. From a finite experience of speech acts and their interdependence with socio-cultural features they develop a general theory of the speaking appropriate in their community, which they employ, like other forms of tacit cultural knowledge (competence), in conducting and interpreting social life. (Hymes, 1973)

It is this sociolinguistic component of the language development process that is often at the mercy of environmental opportunity. The consequences, for a child's educational progress, of differential opportunities to acquire communication skills should be obvious. It almost goes without saying that the acquisition of verbal communication skills by children is vitally necessary to their academic achievement and future social-occupational status in modern American society. Unfortunately, there are large numbers of underachieving school children who have problems in the areas of verbal-expressive skills and verbal-reasoning skills (for a review of the research in this area, see e.g., Grotberg, 1969). This fact highlights the need for developing specialized curricula and teaching methods for diagnosing and prescribing treatment for children with these kinds of educational needs.

In an attempt to address these needs this research project took measurement and training techniques that were developed in the experimental laboratory (McCaffrey, 1970) and tried to adapt them to the needs of elementary schoolchildren and to the situational constraints of the elementary classroom. This gave rise to the two major aims in the study:

1. The first aim was to use a model of communicative competence to develop an interrelated set of measurement instruments which would reliably identify and categorise children at different levels of ability along a scale of competence.

2. The second objective was to take children at a certain level of ability and develop their expressive language skills in a communication training program that was designed and developed for use in the elementary classroom.

3. Educational Outcomes from the Study

To put it most simply, the basic aim of this study was to help children who are poor communicators to become good communicators. The rationale for this is fairly obvious. Underachieving and disadvantaged school children typically have difficulty in school in all of the areas of linguistic competence (Grotberg, 1969; Bullock, 1975). Studies by Heider (1968), Krauss and Rotter (1968), and by Baldwin, McFarlane and Garvey (1971) which show middle-SES children to be more accurate in their verbal communications than low-SES children lend support to the claims of Bernstein (1970) that different social class environments produce children with different communication abilities. While less advantaged children may be as able as others to produce ordinary casual conversation, studies like those just cited suggest that they are less proficient when it comes to the verbal communication of specific referential material — for example, accurate descriptions of pictures or directions for accomplishing tasks. Moreover, children who enter school with difficulties in the areas of verbal expression, verbal comprehension and verbal reasoning also have difficulty learning to read and write in the early grades. By the time their troubles are noticed they are often woefully behind their more competent classmates, and some of them may never acquire basic literacy skills at all. The problem is further compounded by the fact that many elementary school curricula pay surprisingly little attention to the child's own spoken language and not enough speaking and listening exercises or opportunities are provided to sharpen his communicative skills. It is not inconceivable that a reticent child might get through the school year quite well with a minimal production of spoken language. The British government Bullock Report (1975) found the incidence and frequency of pupil-talk in English classrooms to be highly variable and cited numerous instances of classrooms where excessive amounts of teacher-talk was the norm.

A lack of emphasis on the child's own verbal-expressive skills is to be regretted for a number of reasons. While most children enter school with some basic repertoire of pre-acquired oral language skills, there is enormous variability in the extent of this repertoire and in the levels of langauge development reached among different first graders. Developmental assessment of level of language capacities is essential for the planning of appropriate language exercises for individual children. In addition, as Chomsky (1972) and Palermo and Molfese (1972) have recently argued, there is still a good deal of language development left to take place during the elementary school years, and Chomsky (1970) has further suggested that literacy acquisition (especially reading) may be closely interrelated with other aspects of the child's linguistic development. One vital aspect of language development in middle childhood is the child's control over the spoken word. A child who cannot produce in his spoken language well-formed sentences, accurate descriptions, clear directions, may very well be unable to include these things in his written compositions. And he may not understnad them when they are produced by others. As the famous American linguist Edward Sapir pointed out many years ago, "...all voluntary communication of ideas, aside from normal speech, is either a transfer, direct or indirect, from the typical symbolism of language as spoken and heard, or, at the least, involves the intermediary of truly linguistic symbolism." Spoken language is, "by whatever devious ways we follow the process, the historic fountain-head of all speech and of all thinking." (Sapir, 1921, p. 21).

The notion of "communicative competence" that is being used in this study is one way of characterizing the importance of the acquisition of expressive language skills for the child's educational development. By applying our model, therefore, together with its testing and teaching techniques, to a language skills program in a second grade classroom it is anticipated that these experimental procedures could become the basis for curricular innovations aimed at providing the relevant materials and learning opportunities that will foster practical, purposeful verbal communication skills in school children.

PART I

Evaluation of Communicative Competence

PART I

Evaluation of Communicative Competence

In order to discuss the Evaluation goals of the research it will first be necessary to outline the theoretical rationale underlying the particular model of communicative competence to be used in this study.

I. Theory

1. Definitions and Limitations

We begin by defining and limiting the scope of the enquiry. Given the social and practical importance of functional language abilities, we were concerned with developing techniques for measuring such abilities in ways that would be as naturalistic as possible within the normal limits of controlled experimentation. We wanted to observe and measure children's language usage "in situ" i.e., in a functional context where the child would be motivated to engage in spontaneous speaker-listener dialogues with other partners. To achieve this kind of testing environment the evaluation tasks, situations and materials had to have some degree of ecological validity and to attain this we had to try to incorporate real-life constraints into the communication problems the child would be faced with. Too much of what passes for assessment of language abilities takes place in artificial settings where an adult confronts a child with a question about an imaginary problem. This is evaluation in vitro rather than in situ, and it can be criticised on a number of grounds (see, e.g., a recent criticism by Spekman, 1980).

We believe that a key concept in this issue of functional language assessment is representativeness. The testing situation should at least attempt to simulate real life problems and conditions, and the variety of utterance forms that are elicited should be representative of the child's pragmatic repertoire if generalisations are to be made, based on test data, about general competence. This notion of representativeness led us in search of a typical communication problem which had ecological validity and which could be simulated reasonably well in the experimental testing situation. We chose a problem-type which both defined and limited the scope of our inquiry. We chose to work with a situation which involves what Rommetveit (1968) has called the "dog-leg of communication". This is the frequently encountered situation where a speaker has to tell a listener about a non-verbal phenomenon and the listener is required to react to or act upon that phenomenon on the basis of the verbal information received. On the speaker's side, the person begins with a non-verbal event (NV) which he then encodes into a verbal message form (V). The listener begins with the verbal message (v) and uses that information to behave towards the non-verbal phenomenon (NV). The chain or dog-leg thus runs -- NV⟶V/V⟶NV. A good real-life example of this kind of communication problem was provided by one of the teachers we worked with. Suppose A has a sister C who is arriving at the airport tonight but A can't go to meet her. A asks a friend B to drive out to the airport to pick up the sister. B has never seen C, so A describes C's salient characteristics and B carries this information in his head on the way to the airport. The proof of the communicative competence of A and B (as speaker and listener) is the correct identification of C by B.

This kind of paradigm provided us with a frame of reference for tackling the measurement of communicative competence in a delimited but relevant way. However, two other considerations entered into the picture. The kind of competence we were interested in implied being competent in more than one situation and for more than one communicative purpose, so this kind of referential communication paradigm is to be seen as a starting point and not as an end in itself. In addition, our definition of competence had to be relevant not only to our evaluation purposes, but also to the goals of the Training program later in the project. Operating within these constraints we arrived at a very general and very simple working definition of communicative competence, as follows:

> "Communicative competence is the ability to use a variety of language for a variety of purposes in a variety of situations, about a variety of objects and/or events, to a variety of listeners."

The kind of versatility demanded by this definition should probably be taken as an ideal goal for both development and education. But it does make the point that assessment (or training) that is genuinely representative of the competence it seeks to measure (or foster) must deal with a matrix of linguistic, pragmatic, situational, phenomenal and interpersonal variables. In other words, Evaluation of general competence, if it is to be relevant to life as well as to school, cannot be restricted to artificially isolated aspects of the process, but should test an ecologically valid range and variety of skills. Similarly, in the teaching and training of communicative competence the child should be involved in learning which exposes him to a variety of communication experiences.

The exercise of communicative competence as we have defined it obviously implies the activation and interaction of various cognitive and linguistic abilities by the speaker-listener. An analysis of the basic skills involved is presented next in the form of a theoretical model.

2. A Model of Communicative Competence

To construct a model of communicative ability along the lines we have just outlined, we chose to work with the experimental "screen" paradigm that is popular in referential communication research (Glucksberg, Kraus and Higgins, 1975; Glucksberg and Danks, 1975; Dickson and Patterson, 1979). Here two subjects, a speaker (Encoder) and a listener (Decoder) are not visible to one another, but must communicate about a set of stimuli only by verbal means. In the typical experimental situation each partner in the dyad has access to only limited amounts of information about the task and each must depend on the other for the construction and transmission of messages so that appropriate response behavior can be made contingent on the verbal messages. For experimental purposes this encoding-decoding problem approximates the sorts of natural situations we spoke of wherein a speaker has to go from a nonverbal state of affairs to a verbal message and a listener has to go from the message to some nonverbal response or result.

Having chosen an experimental heuristic, we next proceeded to develop a rationale for its operation. Our rationale is derived from a model of communicative competence (MCC) developed by McCaffrey (1970). Though conceived independently, this model deals with communication topics (perceptual accuracy, egocentricity, communication effectiveness, etc.) that have also been treated by other researchers (e.g., Flavell, 1968: Garvey, 1972; Hunt and Kirk, 1973;, Glucksberg, Kraus, and Higgins, 1975). However, the MCC does differ from many of these studies in a number of important respects. Many of the studies of children's encoding-decoding behavior in the past have been open-ended and exploratory, and have usually been designed not to test any particular theory or model of the communication process but, instead, to study assortments of variables piece-meal fashion. A second criticism of previous research is that few, if any, studies have made an adequate linguistic description of the functional speech employed by the encoders in the dyadic communication situations. The field of communication research in general is beset by conceptual and methodological problems (Dickson and Patterson, 1979), and the available models all have major weaknesses. Higgins (1977), for example, criticises the limitations of the Flavell model since "some of its categories subsume other aspects of communicative competence that should be distinguished and independently examined" (1977, p. 137). The state of the art is no better out in the schools where good oral communication programs are conspicuous by their absence and current curricula usually lack both a theory-base and a data-base to justify them.

The model we have constructed, on the other hand, attempts to address some of these needs by going beyond fragmented, a-linguistic descriptions of behavior, in order to formulate an integrated account of constituent psychological and linguistic processes which underlie speaking and listening competency. A similar need is identified by Higgins (1977), who concludes his critique of communicative competence research by arguing that what is needed, "is more information concerning the processes and skills involved in effective communication. To obtain such information a greater variety of communication situations must be studied, and more precise measures of communication must be used" (1977, p. 135).

If we can adequately describe the most characteristic components of communicative competence, we might be able to identify necessary (though not necessarily sufficient) conditions for the acquisition and operation of that competence. The model, in turn, allows us to generate assumptions and testable hypotheses about the interaction of processes during the act of communicating. It also provides a rationale for constructing measurement instruments which can be used not only to evaluate competence, but also to assess the influence of training on competence.

This then is the conceptual strategy of the study, with the aim of producing evaluation and training programs which have both a theoretical and an empirical basis. The model is constituted by the following proposals and assumptions:

a) Instead of attempting an exhaustive analysis of all of the variables that could conceivably be involved in a proper definition of communicative competence, the model instead tries to specify what must be a basic, minimal set of abilities that are possessed by an "optimal communicator" (this strategy is somewhat akin to Chomsky's (1968) attempt to specify the grammatical knowledge that must be possessed by an "ideal speaker-listener"). Thus by starting with an objective conceptualisation of the communication process, the model can then be applied to an analysis of the verbal dialogue behavior of the school-age child.

b) Drawing upon the insights provided by previous research, the model next describes the structure of competence by targeting three basic domains of ability -- perceptual, cognitive and linguistic — which underlie competent, interpersonal verbal communication. It further proposes that these three basic requirements can be applied to both the speaker's and the listener's roles in dyadic situations.

c) The model further assumes that, in the act of creating a verbal message for another person, these component processes are not activated sequentially by the speaker, but rather that they operate simultaneously and in parallel in the decision processes which determine the content of the message (Neisser, 1967).

d) Finally, and most importantly, the model (following Brown, 1965) offers a detailed analysis of the perceptual, cognitive and linguistic conditions necessary for optimal Encoding and Decoding behavior within the scope of our paradigm, as follows.

The Encoding Process

1. The first requirement for an encoder is a very perceptual one. He must be able to make an informational analysis of the stimulus array that is the object of communication, so as to identify the distinctive properties of the target stimulus to be transmitted. Many of the children in the McCaffrey (1970) experiments failed in this aspect of the task because they failed to encode just those critical features of the target that made it distinctive.

2. The second requirement is primarily a cognitive one. The Encoder must realistically assess the informational needs of his Decoder. This involves decision processes, comparative judgments and the "representing to oneself the attitude of the other." The egocentric viewpoint of the young child makes it difficult for him to put himself in another's shoes and this problem of adopting a shared perspective interferes with his communication accuracy (Piaget, 1926; Flavell, et al., 1968).

3. The final requirement is the linguistic one. It is necessary for a speaker to control the lexicon and the grammar of a language in order to construct an adequate verbal message. This may be taken for granted for an adult speaker who has a fully developed linguistic repertoire. However, it is a requirement that is not always fulfilled in the communications of children of 6-8 years who have not yet fully mastered the syntax or vocabulary of their language (Chomsky, 1969, 1970; Palermo and Molfese, 1972). The notion of language "control" here has to be seen in a context of sociolinguistic maturation, for the child's attempts to map his or her communicative intentions onto linguistic forms operate under the constraints of his or her linguistic, cognitive and social development (Bates, 1976).

The MCC proposes that these three conditions constitute the basic "minima" for production of accurate messages. In considering receptive, or decoding ability the MCC again postulates three basic requirements which are the point-for-point analogs of the encoding conditions above, as follows:

The Decoding Process

1. Perception. Just as the encoder translates the discriminated properties of a referent into a verbal message-form, so the decoder in responding to a message has to make a "discrimination in reverse" (Rommetveit, 1968) and go from the words to some internal set of appropriate mental representations and/or to an external stimulus configuration. As Roger Brown (1965) puts it, one has to be able to "cash words into referents."

2. Cognition. If the Encoder also has to assess the informational needs of the "other," what complementary cognitive processes must operate in the Decoder? If the message he receives is over-communicative, he has to be able to handle its redundancy; if it is under-communicative, the Decoder has to know what it is that he has not received, i.e., a knowledge of his own informational requirements. For

example, he has to be able to recognize a mismatch between a message and an array of referents for which the message is intended. He must be able to encode the critical omissions into a message code for himself and/or others. [This can take an informative-directive form, e.g., "You forgot to tell me where to put the ...," or an interrogative one, e.g.,"How many X are there?"] Detection of critical omissions in a message may require a further discrimination of target features by the Decoder which parallels the original perceptual discriminations made by the Encoder.

Common to both Decoder and Encoder in this domain is an act of self-reflection, as each contemplates the informational requirements necessary to complete the task. The kind of role-taking or perspective-taking that ensues from this reflection will, in children 7-8 years, depend partly on the complexity of the task, partly on the child's level of operational thinking and partly on the kinds of linguistic operations that need to be mastered and expressed in the communicative message (Bates, 1976).

3. Linguistic. The analog of the lexical and grammatical control that is necessary for accurate encoding of experience is, of course, receptive mastery of the language by a decoder such that he can easily derive the basic structure and meaning of a perceived sentence. The two tasks are probably almost identical in a linguistic sense, except that the Encoder constructs while the Decoder translates language. The latter's situation has additional psychological aspects, again involving a structured system of referents. The Decoder must have a set of referents for language stored somewhere, in some form, as a basis of the meaning of the words he hears. This system of referents has to be derived from common environmental and cultural experience if it is to have any communicative value whatsoever. And just as the Encoder has to have grammatical control of the words he uses, the Decoder must be able to exercise some control over his system of referents if his response to the verbal message is not to be too specific or too generalized. The level of epistemology that is implied in this kind of shared knowledge might possibly be regarded as a later extension of the "shared presuppositions" which Bates (1976) has claimed are implicit in the child's early pragmatic-linguistic development.

The preceding discussion represents the basic theoretical rationale of the Model of Communicative Competence (MCC) which will be applied in the Evaluation phase of the study. Obviously our description is an oversimplification of the complexities of the communication process, but a simpler model is preferable in the study of the grade-school child whose developmental and experimental considerations further confound the interaction of perceptual cognitive and linguistic processes in oral communication.

Recent applications of the MCC (McCaffrey, 1970; Mandelkorn, 1973; Mandelkorn & Corman, 1973) have been made in both evaluation and teaching studies. In order to provide a frame of reference for these applications (as well as for the present study) we will briefly compare the assumptions and strategies of analysis involved in our model with the research strategies involved in other studies of communication skills in children.

3. Critique of Selected Studies[1]

To make sense of the variety of communication parameters that were encountered in this review of related research we categorised and synopsised from each study the data most relevant for comparison with the present study. The results will be summarised very briefly as follows:

Studies reviewed:

Leler, 1967; Baldwin, 1966; Glucksberg, Krauss & Weisberg, 1966; Moore, 1971; Fry, 1966; 1969; Piaget & Inhelder, 1956; Krauss & Rotter, 1968; Cole, Gay, & Glick, 1969;;O'Donnel, Griffin, & Norris, 1967; Flavell, 1966; Ruth, 1966; Cowan, 1967; Heider, 1968; Koen, 1965; Krauss & Weinheimer, 1964; Lantz & Lenneberg, 1964. (For a more recent, exhaustive review, see Dickson and Patterson, 1979.)

Categories of Analysis:

A. Type of Stimulus Material, and Experimental Situation. The most favored communication set-up in these experiments involved a child-child, or child-experimenter pairing where the Ss were either face-to-face or separated by a screen, and the most popular stimulus materials were toys, pictures, or especially constructed experimental objects. This review led us to conclude that there is a whole range of communication situations and materials that are as yet untapped in experimental research.

B. Types of Response Measures Used. This category of analysis looked for any commonality of response measures used across the communication studies. This time there is a little more agreement between experimenters in that half of them choose to score their subjects' communications by means of overlapping categories like "accuracy" (Heider, 1968; Krauss & Weinheimer, 1964) or "efficiency" (Flavell, 1966; Cowan, 1967) or "adequacy" (Krauss & Rotter, 1968; Cole, Gay, & Glick, 1969) or "effectiveness" (Ruth, 1966). A closer look at the definitions of these measures, however, indicates that experimenters have different things in mind when using the same terms. For instance, "communicative efficiency" is interpreted from the encoder's point of view by Flavell (1966) in terms of the "conciseness" of the message used, whereas for Cowan (1967) an "efficiency" score is, in part, a function of the time (in seconds) it takes encoder and decoder to reach agreement on a response. Similarly with measures like "communicative accuracy" or "adequacy" one experimenter concentrates on the decoder's behavior and scores amount of correct pointing (i.e., selection of a correct target, as, e.g., in Krauss and Rotter, 1968) while in another study (e.g., Krauss & Weinheimer, 1964) the reference phrases used by an encoder are also taken into account in scoring accuracy of messages.

Perhaps one reason why there is a lack of agreement about both the best measures of communicative ability and the definitions of the measures themselves is that few experiments are conceived as part of a broad based research attack on a general problem. Instead, isolated aspects of the communication process are singled out for study and rarely is an attempt made to relate the individual findings to any integrated assessment of the overall development of communicative competence. What is so very often lacking in many of these studies is an objective conceptualization of the basic components of the communication process before the experiments are undertaken.

---

[1] For a more detailed account of this literature review see Appendices A - C_2 in McCaffrey, 1970.

C. Style of Exchange, Participant's Role, Direction of Communication, and Task Characteristics. To provide additional examples of the relatively narrow focus of research in this area and of the neglect of important aspects of communication behavior in recent research, we further categorised studies along dimensions chosen to represent a theoretically possible variety of conditions whose interaction creates potential communication situations. We categorised studies on four basic dimensions of communication -- Style (Monologue vs. Dialogue), Role (Encoder vs. Decoder), Task characteristics, and Mode of communication (e.g., from a Verbal message to a Non-Verbal response, or v.v.). Existing studies can be cross-classified through various permutations of these dimensions, but the intersection of categories also allows us to identify potentially researchable communication behaviors that have not yet been studied.

Our use of this matrix analysis on the studies we reviewed led us to conclude that very few of them attempted to study a representative cross-section of communication behaviors. We found an asymmetric distribution of research interests across the categories which indicated that most experimenters preferred to work within the narrow focus of a carefully controlled monologue situation where the dependent variable frequently was the simple act of "pointing." The study of sustained dialogue was even more infrequent, most likely because of the difficulty of analyzing and measuring the qualitative linguistic data generated in a two-way exchange.

In conclusion, our review indicated that much research into "communicative competence" is characterised by a fragmented assortment of experiments, which frequently seem not to be based on any obvious theoretical rationale. The researcher's choice of communication variables to be studied, along with experimental situations in which to study them, has often been too narrowly conceived so that the generality of the findings is limited, and it is frequently difficult to obtain any notion of the overall communicative competence of the subjects in such studies. Representative assessments are badly needed if we are to make generalisations about competence that are properly qualified as to age of subject, and type of material, task, or situation. The kind of detailed experimentation that is required here could lead to more relative but at the same time more "normative" statements about what a child can do at a given stage of development. A similar argument has recently been proposed by Bates for future research on the relations between pre-operational thinking, role-taking, and ego-centric speech:

> In the future, when role-taking is invoked as an explanatory principle, it will be necessary to specify what kind of role-taking, at what level of complexity, is required to perform a certain linguistic operation.
>
> (Bates, 1976, p. 343)

4. Some Experimental Tests of the Model

Our review of research led us to identify the need for a theoretical paradigm from which experimental research strategies could be logically derived. In an attempt to meet this need we have chosen to work with a model which was initially conceived to provide a rationale for the construction of a representative assortment of measures of communicative competence. These could be used in basic or applied research, a) to measure both receptive and productive competence in the same child, and b) to evaluate the child's strengths and weaknesses in each of the three domains of ability (linguistic, perceptual and cognitive) which we believe are essential to communication performance.

Prior to the present studies some direct applications of the model were made in laboratory and educational settings. Some partial tests of the model's assumptions can also be found in some recent research work and we will deal with these first. Growing interest in and recognition of the utility of the encoding-decoding paradigm as an experimental tool has led to a number of additional studies of children's oral communication in recent years. In contrast to the earlier studies, in this area, a number of these have raised explicit questions about the role of one or more of our three component processes in children's communications. Shantz and Wilson (1972), for example, studied child-child dialogue to measure the influence of egocentricity on communication accuracy among 7 1/2 year old second-graders. Likewise, the dyadic situation was also used by Baldwin, McFarlane and Garvey (1971) to test the influence of socioeconomic factors on the perceptual and linguistic coding of critical attributes of target figures among fifth graders. Hunt and Kirk (1973) also used the communication dyad to measure differences in egocentric communication and representational competence among pairs of Headstart pupils and pairs of regular nursery-school pupils.

Direct applications of the model have also been made by McCaffrey(1970), Mandelkorn (1970) and Mandelkorn and Corman (1973). In the original pilot-testing of the model in laboratory studies, McCaffrey (1970) successfully used the paradigm and a battery of perceptual, cognitive and linguistic measures to assess comparitive levels of communicative competence among middle-class and disadvantaged preschoolers aged four to five years. The same approach was also used effectively in the Mandelkorn studies to assess communication abilities of young retardates.

In teaching/training applications of the model, the basic paradigm has also been used to provide exercises in dyadic communication for children who have scored low on the measures of communicative competence in the test phase (McCaffrey, 1970). Low scorers were paired with high scoring children, playing the roles of speaker and listener in turns, in specially constructed communication tasks that require verbal dialogues for the exchange of descriptions or instructions between the partners. Different kinds of representational materials were used in the tasks -- drawings, pictures, photographs, real objects, jigsaw puzzles, etc. In addition to this kind of "peer teaching," low competence subjects were also paired in the training sessions with adult partners who modeled descriptive or instructional messages that had a pre-arranged syntactic or semantic content. Both the training and verbal modeling techniques were used successfully to improve the communicative competence of disadvantaged pre-schoolers. When compared with a control group of children, the experimentally trained subjects showed improvement on measures of communicative adequancy and egocentricity as well as in the linguistic quality of their verbal messages. Using a similar communication paradigm and related training techniques, Mandelkorn (Mandelkorn 1973, and Mandelkown and Corman 1973) successfully improved performance on a variety of perceptual, cognitive and verbal measures among young retardates. Shantz and Wilson (1972) also successfully used the encoding-decoding paradigm along with "peer teaching" in trios of second-grade subjects to effect significant improvement in the construction of descriptive messages by the trained experimental subjects compared to a control group.

In conclusion, we think that there is sufficient evidence to support the utility of the encoding-decoding paradigm in general and of the model of competence in particular for providing successful research strategies for the assessment and training of communicative competence in young children. Most of the above studies, however, were conducted in experimental or semi-experimental situations. What the Evaluation phase of the present project did was to take these laboratory-tested techniques out into the field to conduct a diagnostic assessment of the expressive language skills of second-grade schoolchildren, prior to engaging these children in a competence training program over the balance of the school year.

II. Practice

1. Sample

Subjects were obtained from a public school system in the metropolitan area outside of the city of Boston, Massachusetts. They were all second grade students drawn from nine classrooms spread across three schools. Approximately 150 students were sampled at the beginning of the school year, but not all of these were suitable or available to participate in the final test sample. Participation of the students had to be obtained through a long political and logistical process that involved negotiations with the school administration, public hearings before the town's School Committee, and the sending of letters of permission home to parents, followed by follow-up letters for non-returns. The sampling and screening process took a whole month before testing could begin. After screening and attrition, 137 Ss completed the Evaluation phase of the project.

Out of these 137, a Core Evaluation Sample (CES) of 97 Ss participated in the subsequent Training program and were re-tested at the end of that program. Within this CES was another Core Training Sample (CTS) of 49 Ss. who were the focus of the communiation training. Since all of the data to be discussed in this report pertain to the CES and CTS we will report here the statistics for N = 97. The sexes were evenly balanced in this group, with 48 boys and 49 girls, and these Ss were also fairly evenly distributed across the three schools. The age range for the CES was 73-107 months, with a mean age of 86.6 months at the beginning of second grade. Socio-economically the children came from a background that was representative of the town as a whole, i.e., predominantly lower and middle class. The town also has a rich ethnic mix with sizeable populations of Italians, Greeks, and Armenians. Many of our Ss came from homes in which more than one language was spoken. In order to equate Ss for basic English proficiency, Ss were eliminated from the original test sample of 137 if English was not the child's first language.

During an orientation period prior to testing, the project staff met with the pupils and teachers in the schools in order to familiarise them with the personnel and purposes of the study. Teachers were interviewed about the language arts programs they used in class, and the pupils' records were examined to obtain additional test data that might usefully be correlated with ours. In order to lay the ground work for good test rapport, the adult experimenters during this period prior to testing also spent time in the various classrooms working as teacher's aides and getting to know the children.

2. Testing

Following the rationale of our model of communicative competence, a battery of perceptual, cognitive and linguistic measures were administered to the total test sample of 137 Ss. Total testing time usually lasted from 60 to 90 minutes and was usually spread over two days for each child. The perceptual battery lasted approximately 15 minutes, the cognitive tests approximately 20 minutes, and the two lengthy verbal-communicative tests together lasted 30-40 minutes. Over 159 hours of testing were conducted and the whole evaluation program took two months to complete.

The details of the Test Procedures will be presented below. The responses to the Perceptual and Cognitive measures were recorded on paper on the spot and scored immediately after administration. The responses to the language tests were tape-recorded, then transcribed later, and scoring was done off of the transcriptions.

The constraints of scheduling and logistics within the different schools did not permit us the luxury of manipulating the order of testing, so the sequence of test administration was fixed. Children usually received the tests in the order -- Perceptual, Cognitive, the first Language test (Action Description), then the second Language test (Picture Description). Any supplementary measures were administered after the main battery was completed. Within the Language tests, children were tested in pairs, so starting order and target sequence were counterbalanced across all tasks. The details of this will be discussed in the appropriate sections below.

(i) Perceptual:

Rationale: To test the child's ability to detect and differentiate distinctive features of target stimuli when the latter are embedded in an array of competing stimuli.

Tasks and Materials:

The perceptual tests were an adapted version of Kagan's Matching Familiar Figures (MFF) Test (Kagan, 1965, 1966; Kagan, Pearson and Welch, 1966). This task requires the child to match Standard Figures (of varying degrees of complexity) against their counterparts when the latter are embedded in arrays of competing figures. Our version was labelled "Spontaneous Search" since the children received no direction or instruction about what to look for (Heider, 1968).

Kagan's figures consist of line drawings of objects, persons or animals. Each test item contains two drawings, the first a standard figure by itself, the second an array of four to six similar figures in which the standard is embedded. We used eight of Kagan's items which, after pilot testing, were arranged in an approximate order of increasing difficulty:

MFF Stimuli

Test Pictures	Kagan Reference #
1) Bottle (Demonstration)	Pre-School level #2
2) Apple (Practice)	Pre-School level #3
3) Flower (Practice)	Pre-School level #9
4) Scissors (Test)	Elementary level #2
5) Tree (Test)	Elementary level #5
6) Lamp (Test)	Elementary level #10
7) Bear (Test)	Elementary level #4
8) Cowboy (Test)	Elementary level #12

Procedure: Testing involved one Demonstration, two Practice, and five Test trials. In each case Ss were shown a Standard Target Figure first and asked to label it. They were then asked to find its exact match in an array of variant figures. On the five test trials, a maximum of three choices was allowed on each trial. If the first two choices were incorrect the child was so informed, but after the third choice, whether right or wrong, no such feedback was given and the experimenter proceeded to the next item.

Recording: All choices were recorded on a standard Test Protocol (Appendix A). In addition an electronic timer was used to record times between choices. The timer was started as each array of figures was being revealed, and E. recorded elapsed times between choices and cumulative times across choices for each test item.

(ii) Cognitive Tests:

Rationale: The cognitive tasks were intended to measure the child's role-taking or perspective-taking ability, and the degree to which this is affected by the child's egocentric viewpoint. In a Piagetian analysis, empathic sharing of a listener's perspective should involve operations of reversibility and decentering which may not be fully available to the child until the stage of concrete operations is reached. Taking into account the informational needs of the listener (which is one of the postulates of our model) involves what we would call an act of "self-reflection" or what Bates (1976) would call acts of "pre-supposing." These acts or pragmatic structures can be encoded onto verbal message forms in different ways and at different levels of complexity. We agree with Bates (1976) that this kind of role-taking is not a static, now-you-have-it-now-you-don't experience, but is a dynamic, evolving process. The articulate expression of shared perspective that the child incorporates into his verbal message form depends on the complexity of the communicative demands which a particular situation imposes on the child at his particular stage of development.

We felt it essential, therefore, to measure our Ss' role-taking abilities with differing tasks which required varied responding under conditions of varying concreteness or abstractness. We took as our starting point Piaget's Mountain Perspectives Task (Piaget and Inhelder, 1956) especially as this has been elaborated by Pinard and Laurendeau (1962, 1970). However, this task involves a rather passive mode of responding from the child and we felt the need to supplement it with similar tasks which would tap a more differentiated range of responding under a greater variety of experimental conditions.

After pilot-testing, the final Cognitive battery consisted of three tasks, House Construction, House Pictures, and Mountain Perspectives, which are roughly arranged along a continuum, from more active to more passive responding. The first task differs from the other two by requiring, not just simple pointing or choosing by the child, but active physical construction of a toy house to conform to a model. The second and third tasks are further differentiated from each other by the kinds of operations and response modes they require of the child. In House Pictures, the child is shown a photo of a perspective of the toy house and has to identify the position it was taken from. Conversely, in Mountain Perspectives he is shown a position in the toy landscape and has to identify the perspective from that position in an array of photos. In the tasks below, House Construction has one test item, House Pictures has four test items, and Mountain Perspectives has three test items.

a) House Construction:

Tasks and Materials: The lay-out of the task is illustrative in Figure 1. The child is confronted with a poster-board map containing two parallel roads, and two house lots, each with a driveway. The materials are small, colored toy blocks (cubes of different sizes), including a small car and tree made from the same material.

Procedures: The task requires the child to construct a house with the blocks to compare with a similar house made by the Experimenter, but the child has to orient his house to the same perspective as the Experimenter's. Success on the test item requires coordination of three critical parameters in the construction -- the locations of the tree and the chimney, and the left-to-right color sequence of the house blocks.

The task is in two parts: first, a Training item (with prompts), then a Test item.

For the Training item, E. begins by constructing a simple house off to the side of the board, using one red and one yellow block, and a smaller red cube for a chimney (see Fig. 1). The Experimenter and child are seated directly opposite each other on either side of the street map. After E.'s house is completed the child is instructed to build a house on his side" so that when you walk down your street and up to your front door, your house looks just the same to you as my house looks to me when I walk down my street and up to my front door." The child is given the blocks only after this instruction is completed. After incorrect constructions, two further prompts are given. The first prompt reiterates the premise of the original instruction. If the child corrects his construction, E. proceeds to the Test item. If the child still thinks he is correct, a second prompt is used to ask the child to consider what is the color of his "front door" compared to that of the Experimenter's. If necessary, similar questioning continues until the child correctly arranges the blocks.

For the Test item, E. constructs a fancier house on the vacant lot on the map (see Fig. 1). The house consists of a fixed sequence of Yellow-Red-Black blocks with the yellow block at the end of the driveway. There is a small chimney block on top of the black block, and the tree is placed behind the yellow block. E. points out the differences between the two driveways and demonstrates how he can drive down his street and into his driveway. He then takes the child's car and drives it down the child's street from the opposite end of the board up to the child's driveway, instructing him as he proceeds: "I want you to build your house so that when you drive in your driveway, your house looks just the same to you as my house looks to me when I drive up my driveway.

Can you do that?" The child is then given the necessary blocks and told to proceed.

If the child's construction is incorrect, E. offers a prompt which reiterates the original instruction in order to verify that the child understands the task. If the child still says the houses are the same, E. proceeds to the next task. If he says they are not, E. asks him to change his house to make it look the same. Then, regardless of the outcome of the change, E. proceeds to the next task.

Recording:

All responses, both correct and incorrect, are sketched and noted for the Training and Test items on a standard score sheet (Appendix B). Scoring is based on this record.

b) House Pictures:

Tasks and Materials: The Experimenter's model house from the first task is used again, together with five Polaroid photos of five different perspectives taken from around the house. Using a clock-dial as a reference point, the child's position, looking at the front of the house, is at 6 o'clock (see Fig. 2). The five views which are presented to him are:

View 'A' (Demonstration) = 6 o'clock
View B = 12 o'clock
View C = 9 o'clock
View D = 2 o'clock
[View E - 3 o'clock - not used]
View F = 10 o'clock

Procedure: The child this time sits in a fixed position in front of the model house and is shown one photo at a time of different perspectives of the house. The task is to point (without changing position) to the location from which he thinks each photo was taken.

The Experimenter removes the child's blocks, and rotates the board so that the child is now sitting looking at the front of the model house. E. explains that he took some pictures of the house which he is now going to show to the child. E. begins with the Demonstration photo of View 'A' explaining that he took it from where the child is sitting. The child's agreement about the perspective is solicited before proceeding. E. then proceeds to display Views B, C, D, and F, one after the other, each time asking the child "to show me where I was standing when I took this picture." The child is required to point to a definite spot on the table top. No prompts are given in this task.

Recording: The location of the child's choices are recorded immediately on the score sheet (Appendix B) whereon the correct locations are indicated by a dot which can be circled for easy recording of correct choices.

c) Mountain Perspectives:

Tasks and Materials: In adapting this version of the standard Piagetian perspective task we followed the construction procedures provided by Pinard and Laurendeau (1970). The landscape is constructed out of three cardboard mountains of differing girth and height (see Fig. 3). The orientation grid and nomenclature of the different target perspectives is taken from Pinard and Laurendeau (1970, p. 314). Five Polaroid photos are used to present the array of views to the child. View 'A' (Demonstration) is taken from the child's perspective and is designated as the 6 o'clock position.

View F (Test #1) is the view from 3 o'clock,
View E (Test #2) is the view from 12 o'clock,
View D (Dummy) is a view from 2 o'clock which was included in the array but was never a test position, and
View C (Test #3) is a view from 9 o'clock.

Procedure: The Experimenter stations himself successively at three different compass locations around the periphery of the mountains. For each of these locations the child is presented with the five photos and is required to pick out the photo that corresponds to what he thinks must be E.'s perspective in each of the target locations.

The child is seated in front of the landscape as illustrated in Fig. 3. E. explains that he has taken some pictures of the mountains, and introduces the Demonstration photo, View 'A'. E. squats down beside the child so that the mountains are at eye-level, and elicits the child's agreement that from that perspective the view from the child's chair and the view in the photo are identical.

E. then spreads out all five photos (including View 'A') in front of the child, walks to position F, squats down to replicate the eye-level view of the target photo, and asks the child: "What did the view look like from right here?" After the child's selection of one photo, E. shuffles the five photos, spreads the array out in front of the child once again, and moves to View E, asking: "Which one did I take when I was standing right here?" The photos are again shuffled after the child's choice, and the same procedure is repeated for the final test item, View C.

The shuffling was not completely random, since E took precautions to ensure that the location of both the Demonstration photo and of the target photo differed from trial to trial.

Recording: After each trial the child's choice was recorded on the standard score sheet (Appendix B) before proceeding to the next item. Incorrect choices were described with reference to the Pinard and Laurendeau orientation grid.

iii) Language Tests

Rationale:

The core of the competence battery is two sets of verbal-communicative tasks which each tap different kinds of expressive language skills. The choice of tasks, materials and language functions was dictated by the requirements of our model that a *representative* variety of language usage be tested in problem situations that are as *naturalistic* as possible (within the constraints of experimentation), where functional language is exercised in a *communicative context*. The latter two requirements were met by our use of the experimental "screen" paradigm and by our testing of *Ss* in pairs in child-child dialogue situations where they were required to communicate about common problems. The requirement that assessment be representative was met by working with two broad categories of everyday language use — language for the description of static objects and events contrasted with language used to explain the passage of events. These two different kinds of language use fulfill very basic functions in everyday communication and, viewed developmentally, they have their origins in the early manifestations of stative and active verb forms in language acquisition (Brown, 1973; deVilliers, 1978).

We evaluated the first kind of language use with a series of Action-Sequence tasks which were specially designed for this study. The other type of descriptive language was measured with a battery of Picture Description tasks which were adapted from communication tests developed by McCaffrey (1970). Both kinds of language assessment used in this study represent evaluations of baseline competence, of what the children can spontaneously do with different kinds of communication problems. The experimenter intervenes as little as possible, dialogue and feedback are encouraged, and both partners can say as much or as little as they want until a particular test item is completed.

Before presenting the details of these two sets of verbal-communicative measures, there is a point of research philosophy which we would like to emphasise. An underlying motivation in this project is a cognitive-developmental approach to child study. As such, we are concerned with the study and evaluation of process rather than product. That is, we are concerned with how children talk and not how they produce right answers. True, the tasks presented to the children have a beginning, a middle and an end, and an obvious goal that is comprehensible to the participants. But our focus is not on some end product such as correct "pointing" or "labeling". Rather, the experimental task is a vehicle that allows us to observe dialogue processes and how children arrive at communicative understanding. What we have to work with are their vetbal messages which are the surface manifestations of the interaction of underlying perceptual, cognitive and linguistic processes.

A. Toy Village/Action-Sequence Tests

Rationale: This communication game requires the child to use explanation and instruction to describe sequential action involving objects which are manipulated in both time and space. Children are tested in pairs. Each sits on either side of a screen and is confronted with a toy village (identical for both). The experimenter non-verbally demonstrates an action sequence (e.g., drive a truck behind a house) in front of the Encoder, who then has to explain verbally what he has witnessed to his partner so that the latter can correctly perform the same action on his side. Built into the pre-arranged action sequences are tests of prepositional usage (in, on, under, between, etc.) developed from a technique suggested by Hunt and Kirk (1973). The children each take turns playing encoder and decoder and each has three discrete action sequences modeled for him by E., with each sequence building on the one that preceded it.

Tasks and Materials: The testing is in two parts: first, a pre-evaluation check on Comprehension of Prepositions, then the Action-Sequence Test. For the comprehension task the children's understanding of the test prepositions was tested using blocks and classroom desks (see Appendix C). For the test of sequential action, a toy village was created using a poster-board map containing roads, on which were arranged a house (without a roof), walls, trees, a bridge, a red block, a yellow block, a toy truck, a toy soldier, and three toy animals — seal, elephant and hippopotamus (see Appendix D).

Procedure: Before evaluating their expressive language, a pre-test was made of the children's receptive understanding of the six prepositions to be used in the main task. Children were tested one at a time in an empty classroom, while E. followed the sequence of instructions outlined in Appendix C. The order of presentation of the six commands was varied for each child.

For the main task, both test partners were introduced to the toy village simultaneously, and shown that the starting arrangement was identical on both sides of the screen. In order to follow the action sequences prescribed in Appendix D, the objects in the village were pre-arranged in a fixed position at the outset. In order to control for vocabulary differences, E. asked each child to name the contents of the village and common labels were agreed upon for various objects. Next, the children were seated on either side of the screen and designated Encoder 1 and Encoder 2 according to a pre-arranged schedule which was designed to control and counterbalance Test Order and Starting Positions across both this and the other Language Tests.

E. was armed with a script (Appendix D) of pre-arranged action sequences, which were to be demonstrated in front of each Encoder. Since each test item began where the previous one left off, the sequence of administration was fixed for each dyad. E. instructed the Ss about the purpose of the game, using the two practice items (Appendix D) to illustrate how the game was played. Each encoder was told to describe as best as he or she could what they had witnessed being done by E., so that the unseen partner could duplicate the sequence of actions on the other side of the screen.

E. began by administering the two practice items in sequence and explained that this was how the game would proceed, with each child alternating roles as Speaker and Listener. Next, the first encoder was given test item 1a and when that had been successfully decoded, the second encoder witnessed and described item 2a, and so on, until each partner had encoded three items and decoded three items. Appendix D also indicates how the six test prepositions were built into each test item in such a way that the required order of use was different for each encoder. That is, each child-speaker was required to use the prepositions in a different order from that in which he had heard them being used as child-listener. Another important part of the procedure was that the Encoder was not allowed to begin talking until after E. had completed the demonstration. The Decoder, on the other hand, was left free to begin moving objects as soon as he felt he had heard enough to start acting upon. However, the Decoders were discouraged from asking questions until the Encoder had finished what he intended to say.

E. played very specific roles in this testing. He was not allowed to ask pointed questions or provide directions, but was only supposed to be a facilitator of the communications who would encourage as much dialogue as was necessary to complete a test item. E.'s specific duties were, whenever necessary:

a) To ask the Encoder to repeat anything that was not heard properly by the decoder;

b) to ask the Encoder to repeat the whole message whenever the Decoder got stymied;

c) in the case of a wrong or incomplete solution, to bring the Encoder around the screen to observe the discrepancy between his message and the Decoder's solution, then to have the Encoder return to his seat to give the necessary additional instructions;

d) to monitor both sides of the board and to ask the Encoder, after the original message, if he wanted to add anything more;

e) to encourage feedback from the Decoder and to ask if the Decoder needs any more information or wants to ask any questions.

f) finally, at the end of an item, E. might step in to tie up any loose ends in the Decoder's solution in order to set the scene for the next action sequence. And, of course, E. always offered praise and encouragement for successful encoding and decoding.

Recording: All dialogues were tape-recorded and E. also made a running record of each test session on a standard score sheet (Appendix E) where any deviations, omissions, incorrect solutions, etc. were noted.

B. Picture Description Tests

Rationale: As we indicated above, the other domain of language use to be evaluated was the ability to describe static objects and events embedded in a matrix of competing information. Three different sets of sub-tasks were constructed for this kind of communication game and they have been adapted from test instruments originally developed by McCaffrey (1970) and later used by Moore (1970). The language elicited this time is description of pictures, with the children again being tested in pairs with the "Screen" paradigm. The Ss take turns describing a target picture embedded in an array of similar pictures, and the array is designed such that the target contains both distinctive and non-distinctive features in relation to the other non-target pictures. The encoder's task is to construct a description of the target that is discriminating enough so that the decoder can correctly identify the target on his side. There are three separate sub-tasks which, in all, present each child with eight test items to encode and eight items to decode. The three tasks are called Production Board, Picture Boxes, and Cartoon Boxes (Appendix F).

One of the advantages of testing the same child under two different communicative conditions was the opportunity it afforded us to observe "competence in action", in the sense that the more competent child who adopted a particular communication strategy in the Toy Village task ought to be able to shift strategies to meet the new demands of the Picture Description task. As Sinclair (1975) has pointed out in similar research with preschoolers, the type of natural, left-to-right, concatenated sentence structure that may be affective for action-sequence descriptions, may be less efficient for conveying information about static pictures or objects. The latter may require a language use that is less orthodox in a grammatical sense, and possibly more idiosyncratic. Another way of saying this is that while syntactic constraints on communication may be greater in Action-Sequence types of tasks, semantic constraints may have higher priority when it comes to constructing messages which differentiate target objects for an unseen listener.

One further distinction should be noted about similarities and differences between the three Picture Description tasks. In the interests of creating a variety of communicative demands on the speakers, some differences in format and content were built into the tasks. The Production Board and Picture Boxes tasks have slightly different formats and procedures, while there are differences in content between these two tasks and Cartoon Boxes. Instead of a static array of discrete items, the Cartoon task has thematic content, since the target picture (or cartoon frame) to be described is embedded in a story line. Part of our interest in including this type of task derives from our experiences in pilot testing this battery and from observations made in the McCaffrey (1970) experiments. We observed that in the picture description tasks a munber of children took an almost "blinkered" approach to the task, concentrating on the target picture to the exclusion of all others. They would then embark on an exhaustive listing of all the details of the target, whether distinctive or not. We labeled this uneconomical style the "taxonomic strategy" and were curious to see if those children would retain or change this strategy when they were shifted to a task where the target picture had an intrinsic relation with the other pictures. The taxonomic strategy creates inefficient communication when applied to Cartoon Boxes. We were interested to see if the more competent, more flexible communicators would change strategies on this task in order to take advantage of the additional, useful information provided by this type of material (e.g., the temporal sequence of the story, or the personal relations between the cartoon characters, etc.).

Tasks and Materials: The pictorial content of the three Picture Description tasks was culled from a variety of popular magazines and comics available at the time of the study. Pictures depicting a variety of real life situations were taken from such magazines as Life, Ebony, Redbook, and Seventeen. Cartoon strips were selected from Sesame Street Comics, and from the Comics section of the Boston Sunday Globe. All pictures and all target stimuli were carefully selected to provide a balanced level of difficulty for both encoders within each test item.

a) Production Board

The "screen" in this task was a sheet of hardboard on either side of which were mounted identical sets of five pictures. A sketch of the array can be seen in Appendix F. Although each child saw the same pictures, he did not see them in the same pattern. In order to control for position effects, the pictures were presented in a different order on either side of this screen. This meant that the target picture was located in a different place for either child. For example, in test Item #1a, the Encoder begins describing a target located in his upper left-hand corner, but on the Decoder's side this picture is located at the bottom right-hand corner. When this latter child next comes to describe his target picture in item 1B, it is located in his upper right-hand corner, while his partner sees it in the lower left-hand corner. Position of targets was counterbalanced in this manner across all the stimulus arrays used in this task.

The pictures were of newly married couples, i.e., five couples for each array. They presented a cross-section of racial and ethnic characteristics, and a variety of clothing styles and physical features. Each target picture, therefore, contained both common and unique features. There were two test items (or "boards"), each in two parts. On Board #1, each child encoded once and decoded once, and similarly for Board #2.

b) Picture Boxes

In this task the screen was still in place but the pictures were mounted on boxes attached to a sheet of poster-board which lay flat on the table in front of the child. Five pictures were again used, and the design of the array can be seen in Appendix F. Each picture was attached to the top of a plastic box which could be flipped up. A minor contingency situation was added to this task by placing a plastic token under the target boxes on both sides of the screen. This identified the target to the encoder, while it also showed the decoder if he had made a successful choice. This change in format from the previous task added a little more of a game element to this task, and allowed the children to "win" tokens. This technique was added as a minor embellishment to this task in order to vary the interest of the tests and to sustain the children's attention. It was intended to be simply a procedural gambit and should not, therefore, be construed in any way as an attempt to manipulate "learning" or utilise "reinforcement" during the experiments.

The Picture Boxes task contained three test items, each in two parts to give both encoders the opportunity to describe targets from the same array in each item. As before, the order of pictures on one side is the reverse of that on the other and, whether encoding or decoding, each partner has to deal with target pictures which appear in different locations in each item.

Item 1 has five pictures of Princess Caroline of Monaco playing with a dog, item 2 shows five pictures of a young boy and girl playing on the floor with five different kinds of toy games, and item 3 has five different shots of a set of quintuplets playing together at home . [The latter were taken from a Redbook magazine feature on Dr. Spock and quintuplets from Washington.]

c) Cartoon Boxes

The format of this task is borrowed from Picture Boxes, with cartoon frames mounted on top of flip-up boxes which are attached to a sheet of poster-board in front of each child (Appendix F). This time, however, because of the nature of the picture content the sequence of pictures and the location of targets are not varied but fixed for both children. Each child sees the same story line in the same left-to-right sequence in which it appeared in its original publication. The only constraints which we added were to leave out redundant frames and any frames which contained speech, but this was really only necessary for the comic strip item #1. This showed four frames from a Mutt & Jeff cartoon in the Sunday papers. Items 2 and 3 both came from Sesame Street Comics, with the second showing five frames from "Bert and Ernie Build a Snowman", and the third presenting the five frames from "Prince Charming meets Rapunzel".

Procedures: A similar procedure was followed for presenting all three tasks which were always administered in a fixed sequence: Production Board -- Picture Boxes -- Cartoon Boxes. Test Order and Starting Positions were carefully controlled and counterbalanced across items and tasks. Each child was designated Encoder 'A' or 'B' at the outset, and kept that position throughout the test. Testing began with Encoder 'A' describing the first target picture on Item 1a in the Production Board task. Each child took turns encoding one picture and decoding one picture from the same array. When the next item was presented, Encoder 'B' went first and this kind of alternation of starting position was followed for each item. As Appendix F indicates, this kind of ABBAABBA, etc., sequence was followed throughout, giving each partner equal opportunities to start first or second.

Before each task, both partners in the dyad were shown both sides of the screen to demonstrate the identical nature of the arrays and to ensure a common agreement about picture content. For Cartoon Boxes the children were instructed to study the story line to make sure they understood it before the test began. Since we were interested in communication ability and not knowledge of vocabulary, E. could supply labels for any pictorial details a child was dubious about during this preliminary orientation period.

Instructions were similar for all three tasks, and consisted of telling the Encoders to describe the appropriate target picture "as best you can so that your partner can find the same picture on his side." In addition, in Picture Boxes, the Decoders were told that they would find a token under the box if they chose correctly, while in Cartoon Boxes, in order to draw the children's attention to the story line, E. also added, after the standard instructions, "now tell him what's happening in this one." Decoders were instructed to listen to the whole message before making their choice. The experimenter's role was similar to that played in the Action-Sequence tests, viz., to monitor the dialogue for the purposes of clarification, asking for repetitions, prompting feedback and questioning, offering encouragement, etc. One difference from the Action-Sequence tests was that, due to the different nature of the Picture Description tests, there was no need to have the Encoder come around to the Decoder's side to observe the effects of his message in the Decoder's solution. In the Picture tasks, if the Deçoder selected the wrong picture, E. so informed him, and then asked him to look again, or ask for more information or a repetition of the message. This procedure was followed until the target was correctly identified.

Finally, for the sake of objectivity, the composition of the dyads changed from the first half to the second half of the Language Battery. Children were not tested in Picture Descriptions with the same partners they had had in the Action-Sequence test dyads.

Recording: All of the dialogues were tape-recorded from start to finish for each of the eight items. In addition, E. used the standard score sheet (Appendix F) to keep a running record of the test session, and to record the number and sequence of target choices made by the Decoders.

3. Data Management

Responses on the Perceptual and Cognitive Tests were scored from the test protocols. The tapes from the Language tests were transcribed, all transcriptions were checked twice for reliability and intelligibility, and then scoring was done from the transcriptions, using an elaborate analysis scheme that was designed during the study (see below). A set of analytic measures were then constructed for the major variables from the three domains of competence, all data were coded, keypunched and put onto computer tape to be analysed by the statistical routines of SPSS and the MANOVA program "NYBMUL". When the pre-training and post-training Evaluations were finally completed, each S in the study had amassed approximately six cards' worth of data.

4. The Scoring Systems and the Major Variables of Analysis

i. Perceptual Measures

The raw data consisted of the Time taken for each choice, the number of Correct Choice the number of Correct First Choices, number of Incorrect Choices, and Total Number of choic across the five test items. [N.B., when a child was still wrong on his third and final choice on an item he was scored as having made four choices on that item.]

From these basic scores were derived the following five measures of Perceptual functioning:

1. ACCURACY = Total number of correct choices out of 5.
2. SPEED = Mean time per item (secs.).
3. SPEED x ACCURACY = Mean time taken (across 5 items) to make a Correct choice.
4. SPEED x INACCURACY = Mean time taken to make Incorrect choice.
5. COMPOSITE ACCURACY RATIO:

$$\text{PACC.} = 100 \times \text{Log}\left[1 + \frac{\text{Total Number Correct out of 5}}{\text{Total Number Correct} + \text{Total Number Incorrect}}\right]$$

This measure was intended to use more of the information in the data in one formula, and since it produces a ratio score a log transform was performed to regularise the scale.

ii. Cognitive Measures:

A three-point scale of egocentricity was constructed to measure the quality of the responses on the perspective-taking tasks.

"1-score" = the lowest level, representing an egocentric viewpoint. In the House Pictures and Mountain Perspectives tasks this consisted of selecting the photo that corresponded to the view from the S's position. In House Construction the child constructs a mirror image of the Experimenter's house, i.e., how it looks from the back.

"2-score" = this is an intermediate or transitional level of functioning, where the child avoids the egocentric perspective, but can't pin down exactly what the correct perspective is. This produces a form of mixed responding across the tasks. In the second and third tasks the child often seems to be operating with an elimination strategy where he narrows the choices down to two or three perspectives, but can't decide which one is correct. In the House Construction task a "2-score" usually results when a child can coordinate some but not all of the initial dimensions in the model. He may get the color sequence of the blocks correct, but place the tree or the chimney on the wrong side, or vice versa.

"3-score" = this is the highest level, representing correct, non-egocentric performance.

These three scores accounted for all of the data, and since there was a total of eight items, the maximum possible score was 24.

From these scores were constructed seven measures:

Task Scores (3) = total score achieved on each of the three tasks, summed across items.

Score Frequency (3) = how many of the eight item scores were 1- 2- or 3-scores.

SUMMARY MEASURE (1) = COGSUM = Highest level of performance reached by a child, summed across all scores and tasks (max. poss. = 24).

iii. Language Measures: A Propositional Analysis Scheme

The scoring of the dialogues in the communication tasks concentrated on the expressive language of the Encoders. The transcripts provided a complete record of the exchanges between the Encoder and Decoder, but since there was such a wealth of expressive language and since the training program was going to focus on developing the language usage skills of the less able Ss, the data analyses concentrated on the messages constructed by the Encoder.

Because we were confronted with hundreds of pages of transcript from the two Language batteries, it was essential to boil down the data to manageable proportions. We did this by taking a pragmatic approach to the analysis of message construction, and by performing a task analysis of our own communication situations. This led us to ask two basic questions: How can we best *identify* and represent the *communicative intent* of the speaker, and how can we relate this intent to what the *listener* needs to know in order to be able to *act* at any given moment in the communication chain? There may be some parallels between our conceptual analysis and the work of Dore (1975, 1976) on the grammaticalisation of the pragmatic intentions of the pre-school child, but there is very little material to draw upon which deals with the linguistic analysis of the communications of the older schoolchild in practical, problem-solving situations like ours. We, therefore, adopted an analysis strategy which would subject the form and content of our *Ss*' messages to various levels of analysis, from more general, molar levels to more specific, molecular ones. The scope and sequence of our analyses required us to identify the major *Message Unit* whose constituent *propositions* would fullfil different *functions* for an active listener, and whose overall *organisation* would contribute to the efficiency and success of the communication.

The details of these categories of analysis will be discussed next.

a) The Units of Analysis

The first task was to apply our analyses to a manageable chunk of the encoder's total message. We did this by concentrating on the early part of an encoder's description before interruptions, questions, etc., began to muddy the water. The first chunk of description is the major message unit which we called "*First Message*" (F.M.) This FM is the initial description offered by the Encoder, which begins after E.'s instructions and ends before the first requests for repetition, or added information, etc. It is usually self-defined and self-contained and easy to identify. It represents the purest, least adulterated example of what the encoder spontaneously intends to offer as his or her contribution to the completion of the task. The end of the FM is usually easy to locate, since it is accompanied by a natural pause as the encoder comes to the end of all that he thinks is necessary to say. Where such a natural ending is obliterated by interruptions and questions from the decoder before the encoder has come to the end of his *intended* description, it is still relatively easy to trace the course of the FM around these "roadblocks" through the subsequent dialogues until the natural conclusion is identified. On the transcripts the boundaries of the FM are bracketed and the rest of the analyses concentrate on what is happening inside the boundaries.

In formal definition, a FM consists of a finite set of propositions addressed to a specific task.

The major unit of analysis within the FM is the *proposition* which is a statement or utterance about an action or a condition which is pertinent to the goal of the task. While linguistic units are usually isomorphous with propositional units, the reverse is not always true (this is particularly so in the language used in the Picture Description tasks) This was of less concern to us, however, since we were more interested in a functional rather than a structural analysis of communicative interaction.

It follows from the above definitions, then, that a "good communication" is a set of propositions which are optimal for a given task.

The boundaries of each proposition are also bracketed within the FM.

b) Qualitative Analysis of First Message Content

The bracketing of the FM and its constituent propositions constitutes the form of the units of analysis, and allows us to specify the length of the FM as the total number of propositions. Next, it becomes necessary to delve deeper into the substance of each proposition in order to determine the function that each unit performs in helping the listener accomplish the task.

Viewed pragmatically, the content of the propositions fell into three qualitative categories. A proposition was scored (+) when it contributed positively and directly to goal-directed activity, or enhanced such activity, e.g., "Put the truck ..." is scored as a positive proposition in the Toy Village task. A proposition is scored (0) when it has neutral value, or is ambiguous and does not lead to positive action but leaves the decoder waiting for more information, e.g., in "(Put the elephant) (in front of the tree) ....." the second proposition is scored (0) since there are two trees, and the Decoder has to wait for more information before making his final placement. A proposition is scored (-) when it actively hinders a correct solution, e.g., a child saying "put the truck...." when he should have said "put the soldier....."

This +/0/- scoring system is performed on the transcripts and is the basis for most of the data used to construct the analysis measures to be discussed below. Some sample transcripts and scores are illustrated in Appendix G.

Not all propositions in a message are functional, but they still have the status of conversational conventions which can be noted and dealt with at another level of analysis. Among such conventions we identified Attention-getting ("Hey, John?"), Scene-setting ("you see your truck?"). Self-editing ("wait a minute!"), and so on. These kinds of prefatory or running comments usually do not deal with actions or conditions, but they are still important in their own right. Such kinds of utterances were bracketed and examined after the propositional analyses at a more detailed, more molecular level of analyses where we looked at more qualitative indices of stylistic, syntactic and semantic usage.

There were some occasions, however, when these kinds of Topic-Comment relations (Bates, 1976) were not always clearly separable in our functional analysis. In these cases it might be decided to record the existence of the proposition within the FM, give it a quality rating of zero, and go on with the rest of the analysis. This had the neutralizing effect of adding quantitatively but not qualitatively to the FM. This was in keeping with a general analysis strategy of not trying to solve all the problems at one level of analysis, but of trying to maintain a differentiated structure of analysis so that problematic responses could always be shunted off from one level to another for closer scrutiny and dissection.

c) First Message (FM) Organizational Structure

So far our discourse analysis has looked at the nature and scope of the FM unit, and at the nature and quality of the propositional units which make it up. But this kind of segmental analysis tells us nothing about how the message as a whole is organized to reflect the speaker's overall intent. At the supra-segmental, or what we have sometimes called the "meta-message" level, it is possible to consider the total FM as a single proposition (out of a number of possible propositions) which represents the general idea or thought that the child is trying to express through various pragmatic and linguistic means. Sinclair (1980) has recently identified three kinds of non-linguistic "knowledge" which can be observed in young children's utterances, and we think we can extend her scheme to include the child's knowledge of the message-qua-message, as something over and above its constituent parts. Regardless of how successfully the various segments are constructed, the message-as-a-whole conveys an intent to a listener just as importantly as correct grammar or vocabulary. There can be cognitive and organisational components operative at these higher levels of message construction, and we felt it important to try to grapple with these factors in our scoring scheme.

Accordingly, two kinds of *Organisation* Score were developed, one for the Action-Sequence test and the other for the Picture Description tasks. The motivation to develop these kinds of measures came from the need to differentiate children who achieved the same level of performance on the +/0/- scale but whose messages differed drastically in the manner in which their propositions were organised and sequenced. For example, in the Toy Village test, two children might both generate 20 propositions in their respective FMs, and both have all 20 of these scored positive. However, an inspection of both protocols would reveal that, while one message flows clearly and concisely and follows the natural flow of the demonstrated action sequence, the other has a disordered appearance, is full of re-starts, editings, self-corrections, and presents details out of sequence. Both messages at the *segmental* level provide the same kind of information but, from the listener's point of view, the second message is the harder to understand and to follow.

A similar situation was encountered in the analysis of the FMs from the Picture Description tests. Here the child was not confronted with segmental action, but had to embed distinctive statements in a flow of description. Again we encountered cases of children who produced the same amount of positive propositions but arranged them differently within the message. One would place the distinctive propositions up close to the beginning of the message, while the other might place them at the end of a string of ambiguous remarks, leaving the listener to wade through a great deal of noise and redundancy before receiving something unique that he could act upon.

Separate Organisation Scores were developed to deal with these problems of message structure and the solutions, though technically different, are conceptually similar for the two Language domains. The details for each test will be presented next.

1) Organization in the Action-Sequence Data

For the Toy Village task two requirements were necessary for adequate encoding: to include all the necessary actions and to do so in the correct sequence in which they had been demonstrated. The Organisation Score (or the "FM Score" as it was often called) was therefore built arount two categories -- Omissions and Sequence. A six-point scale was developed for rating the overall FM on each item for major or minor errors of omission or sequence. In the scale below, + = no error, 0 = minor error, - = major errors. The two ends of the scale, as usual, are firmly anchored, but for the intermediate "grey areas" judgments had to be made about the relative importance of errors of omitted information vs. errors of correct information in the wrong order.

Omission	Sequence		Score
+	+	=	6
0	+	=	5
+	0	=	4
+	-	=	3
-	+	=	2
-	-	=	0

With a maximum score of 6 for each item the highest possible score for the task was 18.

2) Organisation in the Picture Description Data

The nature of these picture tasks and the communicative demands they impose on the encoder create substantially different requirements than those encountered in the Action-Sequence tests. Because of this, a different scale of organisational measurement had to be applied to the transcript data here. The Organisation Score for the picture descriptions has to be understood in the context of the functional quality of the propositions as measured by the basic three-point scale of +/0/-. In the course of classifying positive propositions in the FM with this scale we were able to identify individual propositions which referred to the Target picture in a unique, non-redundant way. The Organisational score (also known as the "Unique Location" Score) was based on the existence and location of such unique propositions in the course of the FM. Regardless of how many unique propositions were generated in the message, the Organisation score was equal to the location of the first unique proposition in the message. If this proposition was the second one in the message, the score was '2', if it was the fifth one in the message the score was '5', and so on. For Ss who produced no unique propositions in their FM, the Organisation score = 1 + Total Number of Propositions in the FM. This tactic meant that the scale ran consistently in the same direction for all Ss with low score "good" and high score "bad". This requirement satisfied the basic rationale underlying the Organisation Score that a message is better organised (from the listener's point of view) if information important for the accomplishment of the task occurs earlier rather than later in the message.

An Organisation Score was recorded for each of the eight items in this battery for each child.

iv. Language Measures: The Major Variables

To recapitulate, the Propositional Analysis Scheme subjects the encoder's speech to more and more refined levels of analysis which allow us to ask different questions of the same sample of communication data. The analysis scheme is based on a scoring strategy which begins at the most general level by identifying the major unit of analysis, viz., the first, self-contained corpus of utterances which best represent the speaker's initial, spontaneous communicative intentions. This "First Message" is then, in turn, broken down into its constituent units, or propositions, which are functionally defined in terms of the goals of the task. Next, the pragmatic quality of these propositional units are evaluated for their effectiveness in helping the listener accomplish the goals of the task.[1] This analysis of the parts of the message is then balanced by an evaluation of the organisational quality of the message as a whole, and we have attempted to argue that meaning and intention are communicated at supra-segmental levels as well as segmental ones.

This three-pronged analysis strategy -- aimed at message structure, quality and organisation -- means that one level of scoring evolves from another. We can't have "unique location" scores if we don't have the +/0/- scores, and we can't have the latter without first identifying the beginning, middle and end of the First Message. This modular type of scoring scheme also means that other levels of analysis can be added on to it in order to ask more specific questions about more detailed aspects of message content.

---

[1] There is still another level of analysis which will be considered later. This would relate the scoring categories to the MCC by subjecting propositions which are rated (+) to a further substantive analysis to determine whether their content is strongly perceptual, or cognitive, or linguistic. This would then allow us to make a more differentiated analysis of the relations between propositional content and message effectiveness.

These three categories of analysis were the building blocks for constructing the major variables of analysis in this Evaluation phase of the study. The basic structural and qualitative information supplied by these scores was used to construct all of the other measures which will be discussed below. We will deal with each of the Language Domains in turn.

a) Action-Sequence

1) Prepositional Usage: this measure was simply the number of target prepositions out of six which were correct.

2) Quantity: this is also sometimes referred to as Message Length or Fluency. It is the total number of propositions (positive and zero and negative) summed across the three items of this task, for each child.

3) Quality: also referred to as "basic propositional quality". It is the number of propositions rated positive (+), summed across the three items for each child.

4) Efficiency I: also referred to as the "simple ratio" measure of efficiency. It is:

(Sum Positive Props ÷ Total Number Props) x 100.

This formula allows us to express the quotient as a "percentage of efficiency", and by dividing by Quantity we can correct for individual differences in message length.

5) Efficiency II: also referred to as the "Ratio Gain Score", it is only used in the post-training analyses. It is:

(Post-Training Efficiency I) minus (Pre-Training Efficiency I).

6) Efficiency III: also known as the "Log Ratio" or the "log transformed" measure of message efficiency. This was an experimental measure which was intended to provide an alternative way of looking at message quality, where the effect of the log transform was to regularise the distribution of scores in case the ratio scores had scaling problems. It is basically the log of Efficiency I, i.e.,

$$100 \times \text{Log} \left[1 + \frac{\text{Quality}}{\text{Quantity}} \right]$$

7) Ambiguity: the total number of propositions rated ambiguous (0) summed across the three items for each child.

8) Inefficiency I: similar to #4 above, this is the simple ratio of the total number of "zero" propositions to the total number of propositions for the task:

Ambiguity ÷ Quantity.

9) Inefficiency II: similar to #6 above, this is the log transformation of the Inefficiency ratio:

$$100 \times \text{Log}\left[1 + \frac{\text{Ambiguity}}{\text{Quantity}}\right]$$

10) Organisation: also referred to as the "FM" score, this is the measure already discussed of the level of message organisation on each of the three test items. Score ranges from 0 - 6 on each item, with a maximum of 18 for the task.

11) Composite "Competence": also referred to as the 'Composite Efficiency' Score, this measure attempts to pull together the structural, quantitative and qualitative information provided by individual measures in order to construct a composite profile of overall ability, expressed as a % competence. The formula is:

$$\frac{\text{Quality} + \text{Organisation}}{\text{Quantity} + K(=18)} \times 100$$

K is a constant equal to the maximum score possible on the Organisation measure and is set so that the product is never >1.

12) Composite Efficiency II: this is another experimental transformation of the basic competence ratio in #11, and is:

$$100 \times \text{Log}\ [1 + (\#11)]$$

13) Composite Competence Gain Score: this was used only in the post-test analyses and is the post-test version of #11 minus its pre-test version.

14) Predicted Quality: this is a regression measure of message quality where the effect of variations in message length on propositional quality is removed by a regression transformation on the basic Quality scores. The formulas are derived from regression analyses (to be discussed later) of the Quality-Quantity relations in the message data.

Pre-test:

Predicted Quality = Quality - [5.3 + (.4 x Quantity)]

Post-test:

Predicted Quality = Quality - [7.5 + (.4 x Quantity)]

This transformation was performed on the Summary Task scores for each child. It is similar to the other ratio and logarithmic transformations described above to the extent that it provides a "corrected" Quality score, independent of length or fluency.

[15) Negative Propositions: although negative propositions (scored " -") were recorded and tallied, they were not used in the data analyses because they represented such a small proportion of the total proposition scores. In the pre-test, propositions rated minus accounted for only 8% of all propositions, and this figure dropped to 3% after training, on average.]

b) Picture Description:

The analysis measures constructed for this language domain were approximately similar to those used to measure encoders' speech in Action Sequence, with one difference. The different nature and structure of the Picture Desdription tasks made the minus score for propositions irrelevant. Accordingly, the content of the propositions were rated only with + or 0. In most of the measures below, scores were constructed at three levels -- for each item, for each sub-task (summed across items) and at a SUMMARY level for the whole test (i.e., summed across all items and all tasks.)

1) Quantity: this is a Summary measure of the total number of propositions produced in the First Message on all eight items. It represents the sum of all the positive (+) and ambiguous (0) propositions produced by a child.

2) Discriminability Index: also referred to as "message discrimination", or the "categorical measure of propositional quality." This measure is unique to this domain and was especially constructed for these particular tests. The score is based on rating each proposition in each FM in each item for the degree to which it discriminates between the five pictures in each stimulus array. The discrimination content of each proposition was rated on a scale of 0 - 4, using an elimination strategy. If a proposition was so vague, ambiguous or redundant that it could refer to all of the pictures, then, since it eliminated none of them, it was rated "0". (E.g., on the Production Board tasks, saying, "it's a man and a woman," when all of the pictures fit that description.) If the proposition was a little less general, and offered a description which could eliminate one picture for the listener, it was rated "1". If it eliminated two or three pictures, it was rated "2" or "3", and if it eliminated all the pictures but the target picture, it received the top rating of "4". A score of 4 indicated a "unique proposition" which could describe only the target picture out of the five in the array.

Based on this scoring system, the Discrimination Index was defined as the highest level of discrimination score reached in the First Message on each item. This was scored regardless of how many times the child reached that level during the message and regardless of where such propositions occurred. That is, if a child had two propositions rated "4" he would receive a discrimination score of "4". If he had a mixture of scores ranging from 0 to 3, he would have a Discriminability Index for that item of "3", i.e., the highest level reached. [N.B. In Cartoon Boxes, Item #1 has only 4 pictures in the array instead of the usual 5. This meant that the top score possible was '3'. To handle this discrepancy, after all the data were scored, the scores for this item were re-coded upwards to make them comparable to the discrimination scores from the other items.]

For each child, a Discrimination Index was calculated for each item. These item scores were then summed to produce Task Scores, and these were summed to create Summary discrimination scores for the whole test.

It should be pointed out here that this Discrimination variable is a conservative measure which categorises propositions and messages in terms of the highest level of performance reached, regardless of the frequency of any and all other levels of scoring. Frequency of score-type is handled by some of the other measures below.

3) Quality: also referred to as the measure of "simple,"" basic", or "raw" quality of message. This is a frequency measure of the number of propositions scored positve (+) across all the items. however, "+" scores have a different meaning here from their connotation in the Action Sequence test. Here they refer to those

propositions rated "4" or "unique" in the Discriminability scoring, but unlike that Index, all instances of "4" within a message are counted. So the Quality measure is the total number of unique propositions produced on all three tasks. It is thus a Summary measure and it is not corrected for differences in message length. It is, however, a very conservative measure of message quality since it excludes all but the optimal level of performance.

[4) Ambiguity; In an attempt to provide scoring-categories comparable to those used for the Action Sequence data, a zero score was also constructed for the Picture Descriptions. A zero proposition was defined as a "non-4 score" and dealt with all the instances of "non-unique" propositions, i.e., all the scores ranging from 0 - 3. It, too, was a frequency score and was not corrected for individual differences in message length. However, it was later supplanted by the ratio measure of Inefficiency (see below) and so was dropped from the analyses.]

5) Efficiency: also referred to as the "ratio" or "transformed" measure of message quality, or "quality corrected for length/fluency/ etc." The measure represents the attempt to take into account variability of message length at the item, task, and summary levels.

Sets of ratio scores were constructed at each level of the test using the child's Quality and Quantity scores. For each child, on each item, a ratio score was created by dividing the total number of positive propositions by the total number of propositions in the message. This provided three groups of Efficiency scores — eight Item scores, three Task scores, and 1 Summary Efficiency score (= a sum of all ratio scores).

6) Inefficiency: also referred to as the "ambiguity ratio score." The rationale and calculations for this score are similar to those used in #5. After all propositions had been scored plus ("4") or zero ("non-4") a ratio score was constructed for each item by dividing the total number of ambiguous propositions by the total number of propositions in the message. As before, Item, Task, and Summary Inefficiency scores were derived.

7) Organisation: This is the measure that was discussed above in section II.4. iii.c.2. It uses the information provided by the previous scores to determine the existence and location of the first unique proposition in each message. Once this location (from 1st to nth) is determined for each item, a Task Organisation score is derived by dividing the sum of the Item scores by the number of items in the task to produce a mean location score for that task. In the case of Production Board the Item scores are divided by 2, while in the other two tasks the divisor is 3.

It should again be noted here that we are dealing with a mean position score for the first unique proposition in a message, regardless of the frequency of such propositions in the total message.

A Summary Organisation score was obtained by summing all of the item scores and dividing by 8 to obtain a mean location score for the whole test.

A word of advice needs to be offered regarding the interpretation of these Organisation scores. Because '0' scores (= absence of unique propositions) were recoded by '15' to keep the direction of the scale consistent, there is room for confusion at the high end of the scale. This is because a high score can now mean two different things, depending on the item involved. On the one hand, a score of, say, '13' may genuinely mean that the average position of the first unique proposition is thirteenth from the beginning of the message. On difficult items, on the other hand, where the Ss produce few unique utterances there will be more '15' scores, so in this instance, a high Summary Organisation score may be telling you more about the absence of unique propositions than it is about their presence and location.

It can be argued in both of these examples that the ultimate judgment of message quality is similar for both types of message, but the distinction needs to be understood nonetheless.

8) Predicted Quality: also referred to as the "transformed" or "regression measure" of message quality. It makes the same correction for the effect of Quantity on overall Quality that was done in the Action-Sequence tests. A regression transformation was performed on the measure of raw Quality, based on an analysis of the Quality/Quantity relationship, as follows:

Pre-test:
Predicted Quality = Sum of Positive Propositions
minus [2.3 + (.2 x Quantity)]

Post-test:
Predicted Quality = Sum of Positive Propositions
minus [3.3 + (.2 x Quantity)]

This is a single, summary score calculated for each child.

v. Reliability of Scoring

Hundreds of pages of transcript were read and re-read, checked and re-checked in the course of deriving all of the Language measures. The Perceptual and Cognitive measures were essentially scored directly from the protocols, in a relatively straightforward and factual way. When the dozen or so scores from these two domains are added to the more than 50 scores generated for the Language domain, each of the 137 Ss who completed the Evaluation Phase had a complement of approximately 65 separate scores. The same 65 measures had to be calculated at the post-test for the 97 Ss who completed the Training Program. This meant that each child who completed all phases of the study ended up by producing 6 cards' worth of data containing a minimum of 135 separate scoring variables, a truly enormous data processing task. [For the Language data alone, a conservative estimate of the total number of propositions scored comes to over 5000 for N=97, and over 7000 for N=133.]

Scoring reliability needed to be determined for the Language Tests data since this was the one area of scoring where questions of judgment and interpretation arose during the application of the elaborate Propositional Analysis Scheme. To learn this scoring system, three raters were especially trained before embarking on independent scoring of the transcripts. After all the transcripts were scored once, reliability checks were run on samples of 20 transcripts at a time because of the bulk of the data. Each transcript contained the dialogue records for both members of a dyad for all of the Language tests. Each rater re-scored the transcripts scored by the other raters. In this way a representative sample of transcripts was checked for reliability, but a point was made of including the transcripts of all of the target Ss who were categorised as "Low Competent" and who were to participate in the Training Program.

The reliability analysis concentrated on scoring agreement for the basic building blocks of the analysis system — the identification of FM and proposition boundaries, the +/0/- and Organisational scores for Action-Sequence, and the Discrimination, +/0 and message Organisation scores for Picture Description. Minimal agreement was defined as two out of the three raters agreeing on a score. After level of agreement was calculated for all of the transcripts, any remaining scoring disagreements were resolved by discussion among the three raters together.

The results of the analysis showed good inter-scorer agreement for both the test batteries in the Language Domain. For the Action Sequence tests the range of scoring agreement on the various measures was 92 - 96% with an average level of 93% agreement. For the post-training evaluation data this level of agreement had risen to between 98% and 99%. For the Picture Description tests the range of scoring agreement was from 80% to 98% on the initial Evaluation measures, with an average level of 92% agreement. In the post-test evaluations the range of agreement narrowed to 92-99%, with 95% agreement on average.

III. Results

Overview:

After all responses from the three domains (Perceptual, Cognitive and Linguistic) had been scored, the data were coded, keypunched and computer analysed with the SPSS analysis routines. As a first step in measuring the parameters of competence, the sample was trichotomised into levels of general ability as defined by the MCC. The reliability and validity of this categorisation was further assessed with a variety of statistical techniques, including Discriminant Function analysis, Canonical Correlation and Univariate Analysis of Variance.

The Language Domain was then subjected to more detailed analysis, and Correlational and Factor Analyses were used to appraise the interrelations and interdependence of the two pragmatic-linguistic functions involved.

Finally, in preparation for the next Training Phase of the study, the Evaluation Phase was concluded with the formation of Experimental Groups cross-categorised by Competency Status and Level of Peer-Modeling. Each Training Group included Low Competent Ss who were the primary focus of the program and, in order to determine the Initial Equivalence of these groups of Ss prior to training, statistical evaluations were performed of all the major variables of analysis in order to establish the pre-training, baseline comparability of performance levels of these Ss.

1. Categorisation of Competency Status (COMPSTAT).

Because of the wealth of data and the large number of scoring variables involved, an initial, rough categorisation of levels of general ability was made using the response data for N=133 (i.e., the original 137 Ss minus 4 Ss who dropped out). A first pass was made over the data in an attempt to divide the sample into three levels of competence -- Low, Moderate and High — based on the scores from all of the major measures in the three domains (perception, cognition and language). A general classification criterion was established to assign a child's performance to one of the three levels. After all scores were ranked from lowest to highest on each measure, percentile ranks were calculated for each range of scores. A child was then judged to be LOW COMPETENT if his scores on a majority of the measures fell below the 25th percentile. A child was judged to be HIGH COMPETENT if his scores fell above the 75th percentile on a majority of the measures. All children whose scores fell between the 25th and 75th percentiles on a majority of the measures were judged to be MODERATE COMPETENT.

The composition of the groups resulting from this trichotomy was as follows. Forty-nine Ss (37% of the total sample of 133) were in the Low group, 60 Ss (45%) were in the Moderate group, and 24 Ss (18%) were in the High group. The three groupings were fairly evenly distributed across the three schools in the project. The relativity of this kind of criterion-referenced classification of communicative competence should be appreciated. It is quite possible that a subject classified as 'Low' in this sample might show performance levels that would entitle him or her to be categorised as Moderate or High in another sample of children tested in another school district which produced a different range of scores.

2. Validating the Trichotomy of Competence

After the initial rough categorisation of the sample by general level of performance a more systematic refinement of the trichotomy of competence was made with statistical analyses of the COMPSTAT divisions for each variable within the three domains. These analyses focussed on the 97 Ss who were selected to participate in the Training phase of the project and who were to be re-tested with the same Evaluation battery at the end of training. The rationale for selecting these 97 Ss will be presented in more detail in later sections of the report below. It suffices for the moment to know that this Core Evaluation Sample consisted of all 49 of the Ss identified as Low Competent, all 24 of the Ss classified as High, and a further sample of 24 of the 60 Ss identified as Moderate Competent. These three groups constitute the COMPSTAT Trichotomy and are referred to in the analyses below as "Lo/Mod/Hi" or "L/M/H."

i. Univariate Analyses

One-way Univariate analyses of variance were performed on six Perceptual variables, and seven Cognitive variables. Because of the additional factor of Test Order or Starting Position (A vs. B) in the two Language Batteries, two-way Anovas were used to test COMPSTAT x POSITION effects for twelve Action-Sequence measures and seven Picture Description measures. The results of the analyses of these 32 variables (as shown in Tables 39-44) can be summarised quite succinctly. On all of the measures the differentiation of COMPSTAT levels is consistently supported by the data, nearly all of the performance trends are in predicted directions, and in over 75% of the comparisons the differences between the levels is statistically significant. The general conclusion to be drawn from these results is that the differences between our three levels of competence are real and valid ones. We will discuss briefly each of the four domains in turn.

a) Perception: In 4 of the 5 analyses shown in Table 39 , the pattern of scoring follows the hierarchical ordering of performance established by the trichotomy, i.e., H>M>L, where the symbol ">" indicates a superior level of performance by one group over another. The one exception to this trend is the measure of Time to Incorrect First Choice (Table 39.4) where there is no such symmetric ordering of results and none of the differences between the groups is statistically significant. Time to Correct First Choice (Table 39.3), on the other hand, does show a clear-cut H>M>L pattern of Speed x Accuracy relations, with significant differences ($p < .05$) between the groups, indicating that the more competent children take more time when making a correct first choice. This result is corroborated by the independent measures of Accuracy (Table 39.1) and Speed (Table 39.2) which show that children who are more able at the level of general competence are, in the specific exercise of perceptual discrimination, significantly slower and more accurate than children categorised as less competent.

The composite measure of Perceptual Accuracy (Table 39.5) is less revealing than the separate measures. The major difference is between the Low competent and the other two groups who are more accurate in general, but this difference is not significant.

b) Cognition: The seven major Cognitive measures fall into three groups of analyses -- by Tasks, by Measures, and by the SUMMARY Measure, COGSUM (Table 40). As the data indicate, the scoring trends are clear-cut and hierarchically ordered throughout these analyses. For the Tasks, the pattern of performance is significantly H>M>L for House Construction ($p < .05$), House Pictures ($p < .004$) and the Piagetian Mountain Perspective task ($p < .006$), indicating that the more competent groups were responding less egocentricall The frequency of score-type (1, 2, or 3) also revealed significant differences between the three levels of ability (Tables 40.4-40.6)For 1-scores and 2-scores the performances were ordered $L > M > H$, showing that the Low competent Ss were producing significantly more egocentric responses ($p < .03$) and that the Low and Moderate groups were between them, producing more transitional "2-scores" ($p < .08$) than the High competent Ss. On the measure of non-egocentricity, the distribution of 3-scores is again ordered $H > M > L$, with the High group showing themselves to be significantly better at perspective-taking ($p < .001$) than less competent Ss. Finally, the same $H > M > L$ pattern is repeated for the summary measure, COGSUM, showing that the trichotomy of competence very significantly differentiates the three groups ($p < .001$, Table40.7)by level of perspective-taking ability.

The Cognition data also allow us to distinguish a hierarchy of difficulty among the three tasks. When task difficulty is defined as a function of level of possible performance achieved, the hardest task proves to be the one-item House Construction task where the average score for the sample as a whole (N = 97) was only 50% of the maximum possible. The next hardest was Mountain Perspectives, where 57% of maximum was achieved, and the easiest was House Photos, where 75% of the maximum score possible was achieved. A similar analysis of score frequency shows that the sample as a whole was more likely to produce indeterminate or egocentric responses. The most frequent score on the three-point scale of egocentricity was "2" (43%), followed by "1" (31%), while the least frequent was the "3-score" (26%), representing correct, non-egocentric responding.

c) Language: Action-Sequence

The results for the Action-Sequence data (Tables 41,42) show very strong and consistent performance trends. The two-way Anovas [COMPSTAT (3) x Test Position (2)] showed significant differences between the groups on the COMPSTAT dimension, but there were no significant differences within or between groups for the effects of Test Position on any of the 11 variables involved.

Looking first at the summary of the raw data (Table 41), quantity of propositions produced is a direct function of competency status ($H > M > L$), as is the number of positive propositions ($H > M > L$). The reverse is true for ambiguous ($L > M > H$) and negative ($L > M > H$) propositions. The univariate analyses of these data, however, show that the significant differences between the three groups are qualitative rather than quantitative ones. The children in all three groups are all comparatively fluent in the sense of how much they say. Where the COMPSTAT differences appear to be is in the quality of what is said and in the fluency with which it is said. The differences in message length (Table 42.2) are not significant, but across all of the more positive qualitative indices there is a consistent $H > M > L$ trend for basic Quality ($p < .001$), Efficiency I and III ($p < .001$), Organisation ($p < .001$), both Composite measures of Competence ($p < .001$), and in Predicted Quality ($p < .001$). For the more negative indices of propositional quality the trend is $L < M < H$, i.e., the less competent Ss are more ambiguous ($p < .003$) and more inefficient ($p < .001$) in their message construction in this domain.

If we disregard COMPSTAT differences, the data in Tables 41,42 also provide important developmental information about general competence for the kind of language usage required by this task. Children of this age level (7 - 8 years) all seem to be fairly fluent when describing a spatio-temporal sequence of events. Each group produced close to 21 propositions for the whole task, or an average of seven propositions per test item. In addition, the quality of message content was relatively high for the total sample -- an average 70% of all propositions were rated plus. For the three groups this quality/quantity proportion ranged from 62% for Low, to 73% for Moderate, and 84% for High. The majority of the remaining propositions were ambiguous ones (22%), while the proportion of negative propositions was less than 10%, which is the reason they were excluded from analysis. One can conclude, from these findings, that most of these children are quite "competent" when dealing with this type of communication situation.

d) Language: Picture Description

The results for the remaining battery of Language tests are a little more complex, but the general pattern described for the other domains (above) is confirmed here also. The two-way Anovas (COMPSTAT x Position) showed significant COMPSTAT differences on most of the measures, but there were also some instances of Position effects and these will be discussed below.

The summary of the raw data (Table 43) show that the more able Ss consistently outperform the rest on all dimensions. The trend is $H > M > L$ for total number of propositions produced and for amount of both unique (+) and non-unique (0) propositions generated. Looking at the major variables individually (Table 44), the differences between the groups in message length or fluency proved not to be significant. As reported for the Action-Sequence data above, the significant COMPSTAT differences appear to occur in qualitative aspects of message formulation rather than in quantitative ones. Although the pattern of performance for the measure of message Quantity was $H > M > L$, there were no significant COMPSTAT differences for this variable (Table 44.1).

By contrast, the categorical measure of message quality (the Discrimination Index, Table 44.2) produced highly significant group differences ($H > M > L$, $p < .001$) supporting the trichotomy of competence. The Anova of this variable also produced a significant Position effect ($p < .007$), showing children in starting position B to have higher scores on the average than their test partners in position A. There was also the suggestion of an interaction effect for COMPSTAT x Position ($p < .09$) but, since both the Moderate and High groups had comparable levels of scoring for both positions, this finding was followed up with additional univariate analyses of within-groups effects. A separate Anova for Moderates and Highs together produced, as expected, no significant effects. However, the earlier significant Position effect was traced to the Low group where a one-way Anova produced a significant result for Position ($B > A$, $p < .01$). This result will be discussed further below in the section dealing with the Core Training Sample ($N = 49$). Further analyses were also made of the data from the three sub-tasks which go to make up the summary measure of Discrimination. These analyses produced the same significant $H > M > L$ scoring trends to corroborate the pattern established by the summary measure.

Another index of message quality was the measure of basic propositional Quality (Table 44.3). Here again the $H > M > L$ pattern of performance was significant ($p < .001$), but there were also significant effects for Test Position ($B > A$, $p < .06$) and the interaction of COMPSTAT x Position ($p < .03$). Further analyses of these results showed neither the Moderate nor the High groups producing consistent effects for Position but, once again, it was the Low group who produced the uneven performance responsible for the superiority of position B over A ($p < .006$, see below).

It should be remembered that the untransformed measure of basic Quality is not corrected for variations in message length. When this correction is made via the measure of message Efficiency (Table 44.4) the COMPSTAT effect for $H > M > L$ remains significant ($p < .067$) but the Position effect drops out. This immediately raises the possibility that the earlier Position effects for untransformed Quality have a basis in quantitative variations in message length. The previous analyses cited above for message Quantity contained no significant effects for Position, but a re-inspection of that data both for the Core

Evaluation Sample ($N = 97$) and the Core Training Sample ($N = 49$) provides supporting evidence for the alternative explanation. These inspections revealed that, while there were no consistent trends for Position among the Moderate and High groups, there was a definite $B > A$ pattern among the Low groups. However, when this trend was lumped together with the data for the other COMPSTAT groups in the main two-way Anova for Quantity Table 44.1) the effect for Position ($p < .137$) failed to reach significance. This alternative interpretation is further reinforced by the results of the analyses of the other variable corrected for length, the regression transformed measure of predicted Quality (Table 44.7). Here, when the effects of Quantity are removed, the COMPSTAT effect still appears ($H > M > L$, $p < .01$), there is no effect for Position, but there is an interaction effect ($p < .06$) which again appears to be largely attributable to the Low group, wherein Encoder B, on the average, scores higher than Encoder A.

An interpretation, therefore, of all these results for the various measures of message Quality appears to be that the observed Position effects are not true indicators of differential item difficulty in this test battery. Rather, they disguise an effect of variations in message length which seems to have its strongest influence on the Ss in the Low group. That is, it is not so much that "B" items are necessarily easier to describe than "A" items and so help those "B" Ss to produce more positive propositions. Instead, what seems to be closer to the truth is that "B" items elicit longer descriptions from the children and this, in turn, increases the latter's chances of producing positive utterances during such longer messages. This effect appears to be peculiar to Low competent Ss and there are two other observations that provide further support for these interpretations of the real meaning of the Position effects. First, the measure of message Organisation (Table 44.6) shows that Low Ss. when they produce any positive propositions at all are more likely to produce them later in their messages than more competent Ss. The second supporting observation comes from the Quantity data for the High group. These most competent Ss produced identical numbers of propositions in Position A ($\bar{X} = 38.92$) and B ($\bar{x} = 38.83$) and, given this common baseline, failed to produce any consistent or significant differences by Position on any of the other qualitative measures.

The converse of message Quality is, of course, message inefficiency, and the measure of propositional Ambiguity (Table 44.5) revealed that the less able Ss were producing less efficient messages. The performance trend is clearly $L > M > H$, but the analyses produced no significant Compstat or Position effects. The final building-block of the scoring system, the measure of message Organisation, once again revealed the hierarchically ordered pattern of performance, justifying the COMPSTAT distinction (Table 44.6, $p < .001$). The order of performance this time is $H < M < L$, meaning that the more able groups had lower location scores, indicating that they were placing their unique propositions closer to the beginning of their messages, thus increasing the efficiency of their communication. There were no Position effects. In addition, for both this measure and the measure of Efficiency (above) analyses of the sub-tasks data corroborated the results for the summary variables.

Before concluding these analyses of the Picture Description data, we can take the same overview of general levels of competence as we did for the Action-Sequence data. The kind of raw data supplied by the Summary Table 43 provides us with developmental information about the ability of children at this age level to handle this kind of language usage. The same data, when broken down to the sub-task level, can tell us about the differential difficulty of the tasks and test items used in the battery.

Ignoring COMPSTAT differences for the moment, the summary data indicate, once again, that all of the Ss appear to be fairly fluent in their use of this kind of descriptive language. Measured by the standard of Quantity alone, the children in all three groups produced, on average, between four and five propositions per test item. When it comes to message Quality and Efficiency, however, the profile of competence painted by the Picture Description data is the inverse of that shown for the Action-Sequence tasks. For the sample as a whole, the ratio of Quality (+) to Ambiguity (0) is approximately 25% to 75% across all tasks. For the three groups separately this ratio ranges from 23/77% for Lows, to 25/75% for Moderates and 27/73% for Highs. In other words, the children are simply less competent in communicating effectively about this type of material, and their messages contain a good deal more ambiguity than they did in the Action-Sequence tasks.

The use of proportion of positive propositions to gauge mastery or competence in a particular pragmatic style can also be used to ascertain the relative degree of difficulty of each task in the Picture Description battery. An examination of the total amount of propositions generated in each task showed that Picture Boxes produced the greatest number by all three groups (41% of the Total for the battery), followed by Cartoon Boxes (32%) and Production Board (27%) (although the latter has one less test item). Picture Boxes also stands out from the other sub-tasks as the hardest task by reason of the small number of positive propositions generated by it. Out of the approximately 3400 propositions generated for the whole battery by all 97 Ss, Picture Boxes produced only 149 or 4% of them. This was followed, in order of decreasing task difficulty by Production Board (9.5%) and Cartoon Boxes (10.8%). It would appear, therefore, that Picture Boxes is the hardest task for both quantitative and qualitative reasons: it generates both the most speech and the most ambiguous messages.

The same token of proportion of positive propositions can also be used to rank the eight test items in order of difficulty. Remembering that items 1 and 2 are Production Board, 3, 4, and 5 are Picture Boxes, and 6, 7, and 8 are Cartoon Boxes the data showed the items to be ranked: 3 (Hardest) - 5 - 4 - 6 - 2 - 7 - 8 - 1 (Easiest).

From a psychometric point of view the Picture Description test battery also shows a good deal of internal consistency. Each COMPSTAT group registered similar degrees of ease or difficulty with the easier and harder items, relative to its own ability level. In addition, looking only at quantity of propositions generated, each item contributes a very similar proportion to its total task score within and between groups, for the range of item-to-total proportions is only from 10 - 16% for all eight items. The best balanced items are to be found in Production Board and Cartoon Boxes, wherein each item contributes an approximately equal proportion of propositions (regardless of their quality) to its task total. The most variable dimension is, of course, message quality, as we have seen above, and the clearest expression of this is to be found in the hardest task, Picture Boxes, where both quantity and quality of propositions vary as a direct function of the differential difficulty of the three test items.

e) Discussion

The univariate analyses have clearly supported and validated the categorisation of abilities across the three domains of competence predicated by the theoretical model. The pattern of results confirms the hierarchical ordering of performance by the three COMPSTAT groups for virtually all of the major variables, and it provides a positive profile of the competent child. Such a child is less hasty and more accurate in the perceptual analysis of target features, and is more capable of taking the perspective of another person. The competent child is also more able to encode perceptual and cognitive features of a communication problem into a discriminating message form for others, and can express his communicative intent in a manner that is better organised and more efficient, is less ambiguous and of higher pragmatic quality from the listener's point of view. A consistent finding from both sets of Language measures is that communicative competence is not simply fluency in production of message units but means, rather, having control over the qualitative content of those units.

The profile of competence painted here is essentially a multivariate one, in keeping with the multi-component model of the communication process on which the test battery was based. Because of the wide-ranging nature of these assessments, it should be appreciated that there are fluctuations and variations in levels of performance on particular variables in certain domains, both within and between groups. To put it another way, the notion of relativity applies here also in the sense that the general levels of ability exhibited in these evaluations are themselves relative to the child's level of development in each of the domains. This is nowhere more noticeable than in the language domain where the childrer are clearly more competent in dealing with the kind of language required by the Action-Sequence tasks than they are at handling the Picture Descriptions. It should be understood, however, that our assessments of the adequacy of Picture Description language are fairly conservative ones since they involve some strict definitions of what constitutes propositional quality. This notwithstanding, it is clear that the children's control of this particular type of language function still has a good deal of developing to do at this age level. The valuable comparative information provided by our evaluation of the two different language functions has important implications for other developmental research in this area.

The evaluation results also point to the value and utility of a multi-faceted, well-rounded approach to language assessment. Our propositional analysis scheme has provided a valuable, cross-sectional analysis of various levels of message construction ability. This has enabled us to describe important structural organisational features of messages as well as their propositional content.

Finally, our various measures of linguistic efficiency, discriminability and organisation have raised interesting questions about quality/quantity relations in the generation of propositions, and about the "perceptual" and "cognitive" content of various propositional forms. These questions will be pursued in further analyses (below) as we attempt to provide further validation of the theoretical model and of the diagnostic value of its assessment techniques.

ii. Discriminant Function Analyses

Following the univariate analyses of the trichotomy of competence, further analyses were conducted to find the best combination of variables for predicting membership in one of the three COMPSTAT levels. This was done for two reasons -- to further test the diagnostic power of the Model of Communicative Competence (MCC), and to find an optimal set of the most discriminating variables in the test battery. The latter result would allow us to concentrate future analyses on a reduced sub-set of the large number of variables in the Evaluation Study, and this would be particularly useful for selecting variables to include in the post-test multivariate analyses of the effects of training. Use of discriminant analysis and classification procedures would also tell us which domain or domains (as represented by sets of discriminating variables) of the MCC were contributing most to the diagnostic power of the competence battery. The theoretical model would predict that the most powerful predictors of COMPSTAT would be a combination of variables that represented all three domains of Perception, Cognition and Language.

Accordingly, four additional analyses were run using the Discriminant Function Analysis program of SPSS (Table 45). The first analysis looked at three domains -- Perception, Cognition, and variables from the Action-Sequence component of the Language domain. the second and third analyses concentrated on the Picture Description component of the Language domain looking, first, at sets of sub-task scores and next, at sets of summary scores across tasks. The final analysis looked at sets of variables from all the domains together. In all four analyses a stepwise solution method was used, with Wilks' Lambda and Change in Rao's V being used to statistically test the significance of the separation of the groups on the particular variables involved.

a) Analyses and Classification

Analysis #1. Eleven variables were entered into the first Discriminant Analysis, two Perceptual, three Cognitive, and six Language measures from the Action-Sequence domain. The Perceptual measures were Total Number Choices across the five test items, and Total Time for the whole task. The Cognitive measures were the two scores from either end of the egocentricity scale (1-score and 3-score) and the summary measure COGSUM. The six Language measures included, Quality, Quantity, simple Efficiency, Organisation, Composite Competence, and a special Summary measure constructed for this analysis, consisting of a simple sum of all the scores provided by the major variables.

The analysis identified four variables to be the most discriminating -- one Perceptual (Total Number Choices), one Cognitive (the 3-score), and two Linguistic (the composite measure of Competence, and the Summary measure). The Wilks and Rao statistics for these variables were all significant well beyond .001. Two Discriminant Functions were derived of which the First was the more important, accounting for 97.65% of the total variance and showing a canonical correlation among the four measures of .78. The statistics showed that considerable discriminating power existed among the measures prior to being absorbed by the first function (Lambda = .376, $X^2(8) = 90.37$, $p < .001$).

The final step in the analysis was Classification of cases as an empirical measure of the success of the derived discriminations. Table 45 shows the classification table for this analysis, comparing predicted group membership with actual group membership. The classification routine was very successful, identifying 76.3% of the cases ($N = 97$) as members of the COMPSTAT groups to which they actually belonged. These variables from the three domains of competence are thus quite powerful discriminators of ability when used in conjunction with one another.

Analysis #2. The second Discriminant analysis used 18 sub-task measures from the Picture Description language domain comprising, for each of the three sub-tasks, the scores for the major variables of analysis -- Discrimination, Quality, Efficiency, Organisation, etc.

Four variables proved to be the most powerful discriminators, two from Picture Boxes and two from Cartoon Boxes. Two of these--the Discrimination score from Picture Boxes and the Organisation score from Cartoon Boxes -- proved to be the most significant discriminators ($p < .001$, and $p < .004$ respectively). Once again two Functions were derived, with only the First Function important. This accounted for 98.2% of the total variance and the canonical correlation among the four measures was .59. All of the discriminating power of the measures is accounted for by the First Function (Lambda = .65, $X^2(8) = 40.16$, $p < .001$), and it is important to note, for future reference, that the two variables which "loaded" highest on this function were the Discrimination (Picture Boxes) and Organisation (Cartoon Boxes) measures.

The Classification of cases (Table 45) by the Picture Description sub-tasks alone is moderately successful, but less accurate than the first analysis using measures from three domains. The percent of cases correctly classified was 61.86%.

Analysis #3. The third Discriminant analysis attempted to determine if discriminatory and classificatory accuracy in the Picture Description domain of language usage could be improved by using summary measures instead of sub-task scores. Accordingly, six summary variables were all evaluated via a direct rather than a stepwise solution method. The six measures (summed across tasks) were Discrimination, Quality, Organisation, Quantity, and two measures of Efficiency. The results showed that only two Functions were derived, with the First Function accounting for 94.9% of the variance. The canonical correlation between the six measures was .57. The discriminating power of the measures was also significant (Lambda = .65, $X^2(12) = 38.98$, $p < .001$). The standardised discriminant function coefficients also showed that the largest contributions to the First Function were made by the two summary scores for message Discriminability and Organisation.

The Classification of cases (Table 45) is a little more successful using the summary scores rather than the sub-task scores. This time 63.92% of the cases were classified correctly by COMPSTAT group.

Analysis #4: The final Discriminant Function analysis combined the variables from Analyses #1 and #2 above to assess the discriminating power of the set of variables taken from all the domains assessed by the COMPSTAT test battery. Of the 29 variables entered in the analysis, five were shown to have the greatest discriminating power via the stepwise solution method — two from Action-Sequence language ("Composite Competence," and the artificial Summary measure), one from Picture Description language (the Discriminability Index from the hardest sub-task, Picture Boxes), one from Cognition (the non-egocentric "3-score"), and one from Perception (Total Number Choices). Two Discriminant Functions were derived, with the First one accounting for 98.2% of the total variance, and the canonical correlation among the five measures being .82. The five variables very significantly discriminated among the three COMPSTAT groups (Lambda = .31, $X^2(10) = 106.6$, $p < .0001$).

In addition to their discrimination power these variables also provided the most solid and most accurate basis for classifying COMPSTAT membership. As Table 45 shows, this cross-section of variables from all the domains was responsible for the extremely accurate categorising of 81.4% of the cases into their real COMPSTAT groups.

b) Discussion:

These analyses successfully fulfilled their dual purpose of identifying an optimal set of the most discriminating measures and of substantiating the profile of competence predicted by the theoretical model. The clear-cut results and their implications can be summarised briefly as follows:

(1) The best discrimination and the most accurate job of predicting COMPSTAT group membership is done by variables which represent all the domains of competence. The next best job of analysis and classification is done when variables from three domains (Perception, Cognition and Action-Sequence language) are involved.

These results support the MCC's characterisation of competence as a multi-component process whose evaluation must be based on a balanced and representative set of measures.

(2) Assessment and classification are less successful when they rely on information from only one domain, viz., the Picture Description language measures. This domain does contain some very discriminating variables but they have less predictive power than a set of variables which represent all the domains of competence.

Nevertheless, the analyses of the Picture Description language variables do help us boil down the large set of measures generated by these tasks to manageable proportions. The results indicate that a limited set of summary measures do a better job of discrimination and prediction than a large variety of sub-task scores. We will thus be able to make a better selection of variables from this domain for further inter-domain and intra-domain domparisons in this study.

(3) All of the Discriminant Analyses taken together show a recurring pattern of similar results regarding the most discriminating and most representative variables, and this has important implications for the selection of variables for the later multivariate analyses of Training effects in this study. In the Perceptual domain it is clear that some measure of "number of choices" ought to be used in further analyses of inter-domain relations. In Cognition, some measure of 'level of non-egocentric responding' is obviously a best bet and correlational analyses (below) will explore the relations of the '3-score' to other Cognitive variables to help determine the final choice of measure to represent this domain. In the domains of language functioning the choice of discriminating measures can also be narrowed down to two or three variables. For the Action-Sequence type of language the most representative variable would appear to be our measure of "Composite Competence" which incorporates information about Quality, Quantity and Organisation of propositional output. Since it is constructed of these building blocks of the Action-Sequence scoring system, it would be an economical and representative measure which would provide important continuity with other analyses and comparisons within and between domains. The results of both sets of analyses of the Picture Description variables make it clear that the most consistently discriminating variables in this domain are the measures of quality of Discrimination and of Organisation of propositional content of messages. Both of these variables should be included in future multivariate analyses.

In the meantime, if one were forced to pick the minimal number of the most discriminating measures based on the present analyses, one would probably choose the "Big Four": Total Number of Choices (Perception), the 3-score (Cognition), Composite Competence (Action-Sequence Language), and the Discrimination score for the Picture Boxes sub-task (Picture Description Language).

(4) Finally, two observations can be addressed specifically to the Language domain. First, the results of the Discriminant Analyses reinforce our contention that the assessment of language functioning requires a multi-trait/multi-measure approach. One cannot base claims about communicative competence on simple measures of simple responses isolated from the larger context of language use. Our results show the power and value of measurement and analysis techniques which attempt to look at different facets and different levels of the rich and complex language data that is produced in functional communicative contexts. We have attempted to devise measures which deal with both the form and the content, the part and the whole of message construction. The discriminating power of such compound, inter-related measurements has been validated for both Action-Sequence and Picture Description by the above analyses.

Secondly, the complementary but independent contributions made to COMPSTAT analyses and classification by variables chosen from each of the two areas of language usage measured by our battery seems to require a re-interpretation of the Language component of our model. As originally conceived, this component was characterised in very general terms as embodying very broad language usage abilities. No attempt was made to further refine or differentiate the kinds of pragmatic-linguistic skills that might, or ought to be involved, from a developmental point of view. Our somewhat monolithic conception of the Language domain was intentionally vague and left the details to be "understood." Now, in the light of the Evaluation results, our two areas of assessment of language competency - Action Sequence vs. Picture Description — seem to betoken the existence and operation of two separate Language functions, so that we are now essentially dealing with four rather than three domains of competence — one Perceptual, one Cognitive and two Pragmatic-Linguistic. Both areas of language usage have required different scoring systems, both have been empirically demonstrated to operate at different levels of competence in children 7 - 8 years of age, and both have thrown up variables which are equally important for the reliable discrimination and classification of levels of communicative competence.

It seems, therefore, as if our original methodological concern that our evaluations be "representative" of more than one kind of utterance form now needs to be reformulated as an explicit conceptual distinction between substantively different pragmatic-linguistic functions. The nature, proper description and interdependence of the two Language functions will be further explored in correlational and factor analyses below.

3. Interrelationships Within and Between Domains

The next set of analyses will pursue some issues raised by the preceding analyses. Further analytic techniques (item analysis, canonical correlation, regression, etc.) will be used to confirm relationships and patterns of responding detected in the Univariate and Discriminant analyses, as well as to continue the work of boiling down the large number of variables to manageable proportions. Particular attention will be paid to the Language domain as we a) explore the relations between propositional quality and quantity, b) attempt to determine the psychometric stability of the large number of test items and measures in the Picture Description battery, and c) choose the best and most representative variables to include in later confirmatory analyses. Finally, interrelations within and between all the domains will be studied via correlational analyses, and Factor Analysis will be used to help us identify the structure and components of the two major language functions that have been identified in previous analyses.

i. Language: Item Analysis and Canonical Correlation

This analysis concentrated on the Picture Description tests since these contained the largest number of tasks and items in the battery. These tests almost constitute a mini-battery on their own, since the original tasks were constructed to provide a variety of communicative demands and to elicit a variety of responding. The question can be raised as to how well the test items and tasks hang together as assessments of "Picture Description Language" and about how independent or interdependent they really are.

A first analysis was made at the item level, working with eight item scores for N = 97 for two of the response categories highlighted by the Discriminant analyses, Discrimination and Organisation. A Reliability analysis of the two sets of scores showed significant amounts of difference between the measures, but this seems to be determined to a large extent by the differential difficulty of the items. The analyses confirmed the Picture Boxes items to be the hardest in the scale, as we have seen before. For these two categories of measurement inter-item relations were rather low (the highest r was .32) and Cronbach's alpha ran between .4 and .5.

In view of these results and in view of the earlier indications that no single measure can reliably capture the complexities of the language data, it was decided to perform further analyses at the task level by grouping together the task scores (summed across items for various categories of analysis. Three sets of task variables were constructed, therefore, using the measures of Discrimination, Quality, Ambiguity, Efficiency and Organisation from Production Board, Picture Boxes and Cartoon Boxes respectively. These sets of task variables were examined two at a time, and Canonical Correlations of the sets were performed in an attempt to determine the maximum amount of relationship between the groups of variables.

Interrelations at this level of task analysis proved to be much stronger and more significant. In the first pair of comparisons, Picture Boxes x Cartoon Boxes, the canonical correlation was a highly significant .67 (Lambda = .44, $X^2(36) = 74.68$, $p < .001$). In the second pair, Picture Boxes x Production Board, the correlation was .58 (Lambda = .49, $X^2(36) = 64.84$, $p < .002$), while in the third comparison, Production Board x Cartoon Boxes, the correlation between the two sets of variables was .54 (Lambda = .55, $X^2(36) = 53.56$, $p < .03$). It is obvious here that the results follow the hierarchical order of difficulty that exists between the three tasks: the strongest pattern of relations is between the hardest task, Picture Boxes, and the others, while somewhat weaker relations exist between the two easier tasks.

To summarise the results of the item-analysis and the Canonical correlation analysis: it is obvious that a set of multi-variable task scores are more powerful and more reliable predictors of test performance than single-measure, single-item scores in this domain of language functioning. The significant inter-task relations augurs well for the psychometric cohesion of the test as a whole. In addition, it appears that the choice of tasks and measures has produced a nicely balanced and complementary set of assessment techniques. While the inter-task relations are strong, there is still a fair amount of variance left to be accounted for, thus indicating that each task still contributes something on its own to the total test score over and above what it shares in common with the other sub-tasks.

These analyses, together with the earlier Discriminant analyses, have served to evaluate this particular domain of language use at the item, task and summary score levels. Taking all the results into account, the conclusion appears to be that the most powerful and reliable variables (and the ones to focus on in later analyses) are the Summary measures which incorporate all of the information from all of the tasks for the various categories of analysis.

ii. Language: Quality/Quantity Relations

Some of the Univariate analyses (above) indicated that there might be a special relation between the number of propositions generated in a First Message and the quality of those propositions. Particularly in the Picture Description language and in the propositions generated by the group of Low competent Ss there was a suggestion of an indirect relation between Quantity and Quality of propositions. It seems feasible to argue that there has to be some minimal amount of propositions generated for there to be any positive (or ambiguous) propositions in the message at all. But beyond this minimum quantity it is open to question whether fluency of production per se automatically ensures an increase in message quality and, if it does, by what degree or proportion. The quantitative analysis showed no significant differences between the COMPSTAT groups in fluency, but the qualitative analyses did show the groups differing significantly on indices of message efficiency. Clearly some children are more skilled than others in packing more information into messages of similar length.

In order to answer some of these questions about Quantity/Quality relations in the two language domains, a Regression analysis was performed on the measures of message length and of basic, untransformed quality from the Action-Sequence and Picture Description data.

The correlations between these four variables is shown in Table 46 below. These show that there is a substantial and significant relation between propositional Quality and Quantity in both domains. It is also interesting to note that while there is a low relationship between fluency in either domain, being able to control propositional quality in one domain does have a bearing on message quality in the other ($r = .34$, $p < .01$). The Regression results shed further light on the nature of these relationships. For Picture Description the multiple R is the same as before (.66) while R^2 indicated that 43% of the variance in Quality is explained by the linear regression on Quantity. The unstandardised regression coefficient further indicated that the predicted score on Quality is 2.28 when Quantity = 0, and that this score increases by .18 units for each unit increase in Quantity. In other words it would take a five-fold increase in Quantity to add one more positive proposition to a message.

Table 46 : Correlation of Quality and Quantity for the two Language Domains (N = 97)

	1	2	3	4
1. Picture Description Quality	--	.66	.34	.21
2. Picture Description Quantity		--	.19	.19
3. Action-Sequence Quality			--	.60
4. Qction-Sequence Quantity				--

[N.B. When $N = 97$, $r = .20$, $p < .05$, and $r = .26$, $p < .01$]

This result corroborates what we have observed previously regarding the difficulty level of the Picture Description tasks, for this 5:1 Quantity/Quality ratio reflects the amount of ambiguity generated in the language used for this type of communication problem. The children were much more competent in the Action-Sequence tasks and this is reflected in the Regression results for this domain. The multiple R is as before (.60) while the R^2 indicated that 36% of the variation in propositional Quality is explained by message length The relationship between increasing length and increasing quality is, however, much more direct this time. The regression coefficient indicated that the predicted Quality score is 5.3 when Quantity = 0, and that this score increases by .45 units of Quality for every unit increase in Quantity. That is, this time it only requires a little more than a two-fold increase in Quantity to add one more positive proposition to this type of message, and this is half the ratio increase required to improve the communicative value of Picture Description language.

There are some obvious reasons why the fluency-quality relation is more direct in the Action-Sequence domain. The task is much more structured and factual, and there are more constraints on the child's liberty to say what he pleases, since he has to tell "what happened" instead of "what's important" as required by Picture Description. This latter task is a tough test of competence since it requires the child to impose his own structure on the problem and is much more open-ended about the type of language that is required for an efficient solution. It will be interesting to see in the correlational analyses below what kinds of interrelationships are generated by these two different domains of pragmatic-linguistic functioning.

iii. Perception: Correlations

The interrelationships between the Perceptual measures of speed and accuracy can be seen in Table 7 . While most of the correlations are significant, many of them are of only low to moderate magnitude. Speed is most strongly correlated with Time to Incorrect Choice suggesting that greater speed is more likely to be associated with more inaccuracy. Accuracy (number of first choices correct) is virtually interchangeable with the Composite measure of Accuracy which does not seem to add much more to the analysis. A more solidly consistent measure is Mean Time to First Correct Choice, which shows good relations with both Speed and Accuracy, as well as with the Composite index of perceptual efficiency. Most of the relationships in Table 7 appear to follow predictions and the general interpretation that appears to be warranted is that more haste means less accuracy, as we have seen before in the COMPSTAT analyses.

iv. Cognition: Correlations

As before we group the Cognitive interrelationships into Task, Measures and Summary measure (COGSUM). Table 15 shows that the correlations between the three Tasks are low (.8 to .26) again emphasizing the relative independence of these three sub-tasks (though it should also be remembered that each task contains a different number of items). The three-point scale of egocentricity shows a much stronger set of relations both within its own measures and with the Task scores. The three measures also show significant inverse relations with each other, especially the 3-score and the 1-score. The results confirm the discrete nature of the three scores indicating that the scale has good discrimination power. The relation between Measures and Tasks follows the hierarchy of difficulty, with the easiest task (House Photos) having strong relations with 3-scores (positively) and 1-score (negatively). The moderate task (Mountain Perspectives) has a strongly negative relation with egocentricity (-.69), while the hardest task (House Construction) seems to polarise the responding (r = +.4 and -.4 for the 3-score and 1-score respectively).

The most representative indication of cognitive perspective-taking ability still seems to be the Summary measure, COGSUM, which is all scores summed across all tasks. COGSUM correlates well with the egocentricity scale (especially the 3-score and 1-score) and shows solid relations with all three of the independent Tasks, with the degree of relationship again conforming to the hierarchical order of task difficulty.

What is particularly heartening about the results from the Cognitive domain is the evidence they provide about external validation of this particular battery of tests. Performance on the objective Piagetian Task, Mountain Perspective, that was incorporated into the battery, is well predicted by the general measure of cognitive functioning (with COGSUM, r = .70) and it also has a nicely balanced relation with the three measures on the egocentricity scale (with 1-score, r = -.69, 2-score, r = .34, 3-score, r = .46).

v. Inter-domain relations

Having looked at the different domains separately, we can now look at the interrelationships across the perceptual, cognitive and linguistic domains of competence. Table 47 shows the intercorrelation matrix of the pretest scores for N = 97 for 13 major variables — two Perceptual, one Cognitive, five Action-Sequence, and five Picture Description. Using the results of the previous analyses, these variables were chosen a) for their independence, b) because they are summary measures which best represent their respective domains, and c) to provide a set of discriminating measures which ought to be sensitive to training effects and can thus be used in the confirmatory analyses to be reported in Part III of this report.

At the most general level of inter-domain relations, Table 47 shows that the relations between the Perceptual and Cognitive components are weak, and that the relations between these two domains and Language are on the low side. Perceptual Accuracy has its strongest relations with the measures of message Organisation and Quality in Action-Sequence, but shows surprisingly little relation with the Picture Description variables. Cognition correlates best with the Organisation measures in both Language domains ($r = .23$, $p < .05$), which makes intuitive sense, and it has a similar relation with the Discriminability Index in Picture Description ($r = .24$).

For the two Language components the within-domain relations are generally stronger than those between, and the strongest correlations in the whole table are to be found within the Picture Description domain of language functioning. This may be partly due to psychometric reasons since this is the largest test in the competence battery with the largest number of items and tasks which have already been shown to hang together well.

Within Action-Sequence domain the inter-variable relations are moderate to strong. There are strong relations of Organisation to Quality ($r = .62$, $p < .01$) and Efficiency ($r = .47$, $p < .01$), which allow us again to repeat the argument that one must capture both structural and substantive features of discourse in assessments of functional language. The Quantity-Quality relation in Action-Sequence is positive and strong ($r = .60$) as we have already seen, but message Efficiency (already corrected for variations in message length) has a negative relation with fluency ($r = -.35$). Since message Organisation correlates moderately well with both Quantity (.38) and Efficiency (.47) it would seem that the competent speaker has to juggle quantitative, qualitative and structural features of message construction in order to achieve and maintain an optimal signal-to-noise ratio in his communications.

With regard to lexical precision in language use it is also worth noting that the vocabulary measure of Prepositional Usage correlates quite well with propositional Quality ($r = .33$) and message Efficiency ($r = .44$) in Action-Sequence, as well as with Organisation ($r = .23$, $p < .05$) in Picture Description.

In the Picture Description domain, the pattern of interrelations is similar in some respects to that described for Action-Sequence. Message Structure correlates very significantly with message Quality ($r = .77$, $p < .001$), Discriminability ($r = .83$, $p < .001$), and Efficiency ($r = .65$, $p < .001$). The untransformed measure of message Quality correlates strongly with Quantity ($r = .66$) but, once again, the transformed measure of message Efficiency, like its counterpart in Action-Sequence, has a negative relation ($r = -.20$) with message length. Here again, as we saw in the Regression analyses earlier, "the longer the better" does not hold true, from the listener's point of view. With a view to selection of important variables for further analysis, the generally strong performance of the Discrimination and Organisation variables is again noteworthy in Picture Description. In this regard it is also worth pointing out that these latter two variables also are involved in some of the stronger interdomain correlations between Action-Sequence and Picture Description. All of the qualitative measures in Action-Sequence correlate from .23 to .33

with Picture Description Organisation, while three of these same variables correlate .24 to .38 with the Picture Description Discriminability Index. Among these results, it is worth noting that the Action-Sequence measure of basic Quality correlates relatively well with its Picture Description counterpart ($r = .34$, $p < .01$). Running counter to the general pattern is the result for the Picture Description Efficiency measure which shows hardly any relation with any of the Action-Sequence measures.

Overall, the intra-domain relations in the language area provide the most promising results since so many of the correlations are in the predicted direction. The magnitude of the correlations is not high between the two language functions and we would presume this indicates a substantial amount of independence between the two domains. This conclusion will be explored further in the factor analyses below. The limited interdependence of these two pragmatic-linguistic competencies might be traceable to differences not only in task structure but also in the kinds of communicative demands imposed by the tasks which require different types of language usage.

4. Identifying Two Major Language Functions

i. Factor Analyses

The final stage of evaluating the COMPSTAT results had the primary goal of testing our assumptions concerning the functional division of pragmatic-linguistic abilities within the Language domain of the model. A secondary goal was to explore further the patterning of relationships among the variables from all the domains with a view to further reduction of the data. Both of these goals were pursued with the aid of factor analytic techniques.

Taking the secondary goal first, a series of exploratory factor analyses were conducted with various assortments of variables from all the domains of competence, and the results of these analyses will be summarised very briefly here. The number of variables entered in the analyses ranged anywhere from 8 to 46 depending on how many Action-Sequence and Picture Description measures were included. In the case of the latter domain, numerous groupings of both item and task measures were used in the different analyses. The results of these factor analyses were all similar, consistently producing clear-cut perceptual, cognitive and linguistic factors. The number of factors analysed ranged from three to seven with the number of Picture Description factors varying as a direct function of the number of item and/or task variables entered from this domain. These exploratory analyses essentially carried on and confirmed the work of data reduction performed earlier in the Discriminant and Canonical Correlation analyses. The results helped establish the existence of a multi-dimensional factor structure consistent with the structure of the three domains (perceptual, cognitive, and linguistic) delineated by the theoretical model, and the continued differentiation of separate factors for the Action-Sequence and Picture Description tasks laid the groundwork for the next stage of confirmatory analysis.

The results of the exploratory factor analyses helped confirm our choice of the langua measures to be entered into the confirmatory factor analyses which would test our observation that the Language domain appears to contain two relatively independent sub-domains of language usage, each with a separate character and structure. Four of the most discriminating variables were chosen from each language category. From Action-Sequence, these were the vocabulary measure of Accuracy of Prepositional Usage, the log transformed Efficiency ratio, the measure of message Organisation, and the Quantity measure. From Picture Description the four measures were, likewise, Quantity, Organisation, the Discriminability

Index, and Efficiency. On these eight variables a principal components analysis was performed, with varimax and oblique rotations.

The results of these analyses can be seen in Table 49 . Four factors were identified with eigenvalues greater than 1.0, and accounting for 84.8% of the total variance. In the analyses of these factors, the varimax rotation produced the Factor Matrix that can be seen in Table 49 . The solution appears to be clear-cut and straightforward. There are indeed separate factors for both types of language usage and, within each domain, the factors are further differentiated along a Quality-Quantity dimension. The First Factor loads heavily on the three qualitative measures from Picture Description, the Second Factor loads similarly on the three qualitative measures from Action-Sequence, the Third Factor has a very strong loading on Action-Sequence Quantity, and the Fourth Factor likewise has a very strong Quantity loading for Picture Description. In addition the third and fourth "Quantity" factors also show a negative loading on their respective measures of message Efficiency, further confirming the inverse relation between this particular index of message quality and message length. An oblique rotation of the factor structure also produced the same four factors, with the exceptions that the Third and Fourth factors change places and that the relation of Efficiency to Quantity is more strongly inverted in the factor pattern (see Table 49).

ii. Discussion.

It seems clear, therefore, that the two domains of language usage are independent, but the structure of each domain is similar in that each comprises a quality factor and a quantity factor. We decided to use more functional labels to differentiate these two domains in the light of these results. We re-named "Action-Sequence" as Logical Language usage (LL), containing LL Quantity and LL Quality factors, and "Picture Description" as Spatio-Temporal Language usage (STL), with STL Quantity and STL Quality factors. Our choice of these labels is based on our interpretation of some research by Sinclair (1975) on communicative competence in pre-schoolers. Within the context of the kind of communicative problems we studied, "Logical Language" can be defined as language used for logical categorisation, discrimination, differentiation, and classification of static objects and events. Its use would seem to require some control of the cognitive principles of class inclusion and exclusion, and, looking ahead to the training program, we would speculate that practice in the use of LL might have some cognitive benefits. The other kind of language use, STL, seems more concrete and reality-based since it is the language used for the accurate description of the passage of objects and events across space and time. The pragmatic forms required by STL would seem to conform more closely to the child's own natural language, since, in English, such descriptions would usually follow a left-to-right, subject-verb-object sequence. Logical Language, on the other hand, appears to be less "Logical" and less orthodox in its production constraints. Our experience in this study, like Sinclair's (1975), has been that LL discourse need not be in complete sentences nor follow predictable grammatical sequences when used by children 7-8 years of age. The utterances can be fragmented and disjointed, and may contain a good many idiosyncratic propositions.

From a developmental point of view, STL usage ought to be mapped more easily (and earlier) onto the child's perceptual and cognitive experience of the world and, if so, mastery of STL ought to follow a faster developmental timetable than LL usage. This latter conjecture is certainly confirmed empirically by our COMPSTAT results which show markedly different levels of overall competence for the two language functions. Mastery of LL would seem to be more difficult for the child and take longer to acquire. Its possible dependence on cognitive or intellectual functioning would also make it more susceptible to experiential factors, including educational opportunity. If this is indeed the case, then our two language domains are differentiated not only by function but also by how they are learned. If STL and LL differ in their susceptibility to the influences of development or education then there are some specific implications for the effects of training that can be predicted as the outcomes of this study. But there are also general implications for the wider effects of schooling and training on such functional language competencies. Our characterisation of the acquisition of Logical Language functions and the distinctions we have been making between these and STL development have been addressed directly in a recent report on the relations between education and cognitive development. Summarising their cross-cultural research, Sharp, Cole and Lave (1979) stated:

> Education exerts its strongest influence on tasks where taxonomic principles are the criterion of correct classification, where correct performance requires the subject to provide structure to the nominal stimulus array, where problems must be treated hypothetically without recourse to real-world knowledge, or where problem contents are clearly school related... Insofar as functional criteria and real-world knowledge can produce a correct response or the task is well structured, education-related effects are reduced, and age-related effects are more prominent.

The involvement of functional language competencies in these kinds of learning processes should be the subject of future study.

5. External Validation of the COMPSTAT Scale

We conclude this discussion of the COMPSTAT results by drawing attention to some external sources of validation for the categorisation of levels of communication ability achieved by the COMPSTAT test battery and its measures. These sources include independent tests which were incorporated into the battery, as well as diagnoses of ability which were made for special programs in the school system independently of this research program.

In the Perceptual domain the trichotomy of competence was supported by the results of the children's performance on Kagan's MFF test of perceptual discrimination. The order of item difficulty encountered by our Ss was close to that established by the Kagan scale, while level of overall success on the Task was directly related to our levels of competence on most of the measures.

In the Cognitive domain, the independent Piagetian test of Mountain Perspectives was the reference point for the other measures that were constructed for the Cognitive battery. The univariate and correlational analyses showed that performance on this test is strongly related (r = .70) both to degree of egocentricity on the Cognitive scale of perspective-taking and to general level of communicative competence as measured by the whole COMPSTAT battery. The results of the Mountain Perspectives task thus provide some objective validation of the Cognitive domain of our model in particular, and of the scale of comptence in general.

There are three other sources of data from children's records which provide support for our experimental assessments of competence. These come from records of participation in Title I and Remedial Reading programs, and from scores on the Stanford-Binet Vocabulary test. An analysis of the Title I assignments for underachievers revealed that 13 such children were in our sample and, of these, 10 (77%) were classified by us as "Low Competent" in oral language abilities. An additional 22 of our Ss had been assigned to Remedial Readin classes and, among these, 14 (64%) had been classified by our COMPSTAT battery as Low, 5 (23%) as Moderate, and 3 (14%) as High. It would appear to us that these records not only support the effectiveness of our test battery for diagnosing learning ability but they also provide sufficient grounds for speculating about relations between reading ability and oral language development.

The Vocabulary scores for our Ss also allowed us to classify these by COMPSTAT levels. The results conformed directly and significantly to the trichotomy of ability. For this age level, the Low Ss had a mean score on the Binet scale of 9.0, the Moderates had 10.29, and the Highs had 11.04. The between-group differences were significant ($F(2,94) = 7.29$, $p < .001$). Some exploratory analyses of the relation between these Stanford-Binet Vocabulary scores and our Language domain measures provided some interesting results. Vocabulary correlated most strongly with the STL measure of lexical precision, the Prepositional Usage score ($r = .34$, $p < .01$). Its correlations within the LL scale were -- Discrimination ($r = .25$), propositional Quality ($r = .28$) and message Organisation ($r = .27$).

6. Summary Discussion of the Evaluation Results

On the whole, our Model of Communicative Competence has fared fairly well in this application of it to applied educational research. Like most excursions from the laboratory to the classroom the experience has been educational for all concerned and the effect has created a two-way channel of communication between theory and practice. The theoretical model has been extremely effective in generating an evaluation rationale and a set of test instruments capable of reliably and validly diagnosing levels of communicative competence in young schoolchildren at the "age of reason." Diagnosis without prescription is an empty exercise so, as will be seen in the last section of this part of the report below, this diagnostic information will provide the basis for a language intervention program aimed at improving the expressive language skills of the children who were diagnosed as Low competent.

The results of our various statistical analyses of the Evaluation data have, on the other hand, led us to revise our conceptions of the roles and operations of the different domains in the theoretical model. The Perceptual domain, for instance, has proved to be weaker than expected in its relations with the other domains of competence, even though it contributes effectively to the differentiation of COMPSTAT levels overall, as predicted. It is not clear at this stage if the lack of inter-domain relations is due to the choice of the summary measures which represented Perception in the analyses, and some more light can be shed on this by the post-training analyses. The Cognitive component seems to have better predictive relations with the Language domains and we would entertain the hypothesis that participation in the oral communication training program should have positive effects on cognitive perspective-taking ability.

The major conceptual re-working of the model has, of course, had to take place in the Language domain where the somewhat monolithic status of this component has had to undergo refinement and differentiation. The identification of two separate language functions in our data may not seem too surprising in the light of our attempts to build a representative set of test instruments which would test a variety of responding under a variety of conditions. However, the way in which our various analytic techniques (discriminant, correlation factor analyses, etc.) consistently distinguished and grouped the wide variety of language measures on the basis of intrinsic differences of function makes a substantive issue out of what began as a methodological concern. Theoretically, the division of our "Language" domain into what we have called STL and LL components makes sense and it readily links the results of this study with other developmental research in psycholinguistics and in the pragmatics of language acquisition. Within the model itself, the fact that STL and LL have stronger intra-domain relations than they have with the other components of competence is probably as it should be. It was never intended that all the domains of competence in the model be equally balanced and, from a developmental perspective, it would be naieve to expect that a child at this particular age should show an even profile of mastery of all the components of communicative ability. In addition, from a measurement perspective, the strength of the intra-domain relations within STL and LL can also be considered as a reflection of the psychometric stability of the various language measures and as a proof of the discriminating power of the summary variables which represent the respective domains. With regard to the Intervention program, it remains to be seen if and how the effects of training might shift the pattern of the interdomain relations.

We have discussed in detail the implications of the results of the various analyses in the discussion sections above. We will conclude our discussion of the Evaluation program by reiterating our conclusions about some of the more general implications of this study's findings. These deal with three areas -- Measurement of functional language abilities, Development of communicative competence, and the Teaching of Oral Language as an integral part of Language Arts programs.

The measurement issue has to do with both the context of the evaluations and the scope and depth of the measurement techniques themselves. We argued at the beginning of this report that assessment of functional language abilities ought to be done "in situ", and to take place in functional contexts where communication problems that are as realistic as possible should be presented in test situations which are controlled but also as naturalistic as possible. The rich variety of pragmatic language data that was obtained with such an approach in this study testifies to the utility and necessity of such an attempt. The crux of the matter is the representativeness of the assessments obtained. For a complex process like communicative competence the single-measure, single-test approach will not suffice if reliable and honest generalizations about development and learning are to be based on the data. The "multi-measure" approach helps keep researchers honest by forcing them to develop "multi-level" analysis strategies to deal with the "multi-traits" that are revealed in the rich and complex language data produced by children under such experimental conditions. Our Propositional Analysis Scheme was developed as a response to these exigencies, with some degree of success, and it has taught us the necessity and value of constructing measures which capture interdependent facets of the discourse process -- the quality as well as the quantity of propositional units, the structural form as well as the substantive content of messages, and so on.

Developmental issues are not unrelated to these questions of measurement adequancy. Our results have clearly demonstrated that our Ss are much more competent in effective STL usage than in LL usage. This result is partly a function of task constraints, but the tasks themselves embody distinctly different communicative demands which impinge on the child's pragmatic-linguistic and cognitive processes in very different ways. And the development of these processes, in turn, is affected in different ways by age and education. We have discussed these differential influences of maturation vs. schooling in the context of the cross-culture research by Sharp, Cole and Lave (1979). It is unfortunate, therefore, that the bulk of the popular research on the development of referential communication (e.g., Glucksberg, Kraus and Higgins, 1975) has myopically concentrated on a limited set of Logical Language functions, which are harder to acquire and later to develop. In doing so, such research has grossly underestimated and under-represented the communicative abilities of the pre-school and school-age child. Reliable assessment of communicative competence in the developing child must involve evaluation of more than one pragmatic-linguistic function.

The Developmental issue, in turn, leads into the implications of this study for the Teaching of Oral Language in the classroom. A number of children in our sample who were diagnosed as Low competent also turned out to have reading and learning disabilities, as identified by the school system. This reminds us of the contention of the British Bullock Report (1975) that "learning disabilities are language disabilities." But again the need for a differentiated perspective is called for. If the teacher knows that a child is more competent in STL functions she can build upon these oral language strengths when selecting material for reading and writing practice. On the other hand, the direction of influence can be reversed by treating the weakness in the oral component of Logical Languag (LL) functions by exposing the child to examples of such usage in reading and writing material first, prior to speaking and listening exercises in this kind of communication. Ultimately one would expect that the goal of such an integrated Language Arts approach would be mastery of more complex discourse via a synthesis of both STL and LL functions. This would, eventually, benefit the acquisition of literacy in all its forms -- in reading, writing, speaking and listening.

IV. From EVALUATION to INTERVENTION

The Evaluation Phase of the study provided the diagnosis and classification of communicative abilities in our sample of second graders. This information was then used as the foundation of the next experimental phase of this project, viz., the development and implementation of a Communication Training Program which would foster the speaking and listening abilities of the group of 49 Ss diagnosed as Low competent. To prepare for the next Intervention phase of the study (Part II) the COMPSTAT results were used to achieve two purposes. First, those Ss identified as Low competent were divided up into Experimental (Treatment) groups and a Control (NO Treatment) group. Next, there were added to the Treatment groups additional Ss who were selected from the Moderate and High competent categories and who were to play the role of peer models. The Low Ss would interact with these more competent peers during the training program in order to be exposed to a better quality of functional language usage. Before embarking on the training program, statistical analyses were performed of the scores of the experimental Ss on all the major variables of analysis in order to ensure that pre-training starting levels were similar for all groups. Lastly, a set of experimental contrasts and predictions were specified for testing the outcomes of the training program.

We will conclude Part I of this report by describing briefly the details and results of these preparatory stages for the Training Program.

1. Selecting the Experimental Groups (EXPGROUPS) and Peer Models

In order to involve as many of the COMPSTAT groups as possible in the Training Program all of the Ss from the Low and High categories were involved, along with a matching sample of Ss from the Moderate competent group. The breakdown of the experimental design for Training can be seen in Figure 4.

Figure 4. Experimental Design for Training: EXPGROUPS (4) x COMPSTATS (3).

		EXPGROUP				
		I	II	III	IV (Control)	Total
COMPSTAT	L	13	12	11	13	49
	M	--	12	--	12	24
	H	--	--	12	12	24
					N =	97

The primary design of the training experiment is the first row of that figure. That is, the basis of the study is essentially a one-way ANOVA design which compares the performance of 49 Low competent Ss assigned to one of four Experimental Groupings (EXPGROUPS). The composition of these EXPGROUPS was designed to assess the effects of the two major features of the training program -- Practice and Modeling. EXPGROUPS I-III are the treatment groups and the benefits of communication training will be evaluated by comparing the performance of these Ss with that of the Ss in EXPGROUP IV, the no-treatment Control group. A further within-groups comparison will then be made of the three treatment groups to assess the differential effects of peer-modeling within each of these three EXPGROUPS.

The rationale behind these EXPGROUP divisions can best be understood by referring to Figure 5, which relates EXPGROUP status to the COMPSTAT composition of the members of of the groups.

Figure 5. Peer Modeling and Practice in the 4 Treatment Groups

EXPGROUPS	Composition of Groups	Adult Modeling	Peer Modeling	Communication Practice
I	Low/Low	+	-	+
II	Low/Moderate	+	+	+
III	Low/High	+	+	+
IV	Control Lows	-	-	-

The presence (+) or absence (-) of modeling and practice are indicated for each row of Figure 5. This indicates that the major difference between the EXPGROUPS is in the amount of communication practice received (Control group IV gets none), and in the amount and quality of verbal modeling to which the Low Ss are exposed across the three treatment groups. The modeling is of two sorts, adult and peer. EXPGROUPS I-III all receive a constant and equal amount of modeling during the training sessions by an adult experimenter who participates as a partner in the communication exercises. This is to ensure that even Group I obtains some minimal amount of benefit from the modeling of effective language usage. Peer modeling, however, varies across the three treatment groups, in order to assess both the benefits and utility for regular classroom practice of having different types and levels of peer models for less able language learners. Group I (n = 13) is designated as "Low/Low" to indicate that these Low competent Ss will be interacting with each other in the communication training exercises. If the reader prefers, he can think of Group I as a "zero modeling" group wherein the same Ss are both the initiators and recipients of verbal modeling. Group II (n = 12) is designated Low/Mod to indicate that these Low Ss will be working with a dozen Mod competent partners (see Fig. 4). The highest level of peer modeling is encountered in Group III (n = 11), the Lo/Hi group, where these Low Ss will be paired with a dozen High competent Ss. (see Fig. 4) during training. Group IV (n = 13), Low Control, contains the Low competent Ss who receive neither training nor peer-modeling.

It should be noted that the experimental design in Fig. 5 does not manipulate modeling along without practice. The small amount of available empirical evidence suggests that the utility of such a treatment is not at all clear-cut (Cazden, 1965; Brown, 1973). On the other hand, studies that have combined some kind of verbal modeling together with communication practice have demonstrated the effectiveness of this combination of treatments for improving communication skills in children. (Fry, 1966, 1969; McCaffrey, 1970; Moore, 1971, Shantz and Wilson, 1972.) In addition, comprehensive reviews by Cazden (1972), Bates (1975), and Cooper et. al. (in press) have cited numerous instances of the positive effects of peer interaction on cognitive and linguistic learning.

There is a secondary experimental design nested within the overall research design in Figure 4. This is a 2 x 2 design involving groups of 24 Moderate and 24 High Ss, each divided into experimental and control groups. That is, for each dozen Moderate or High Ss who participated in the training sessions, matching samples of Moderate and High Control Ss were selected. This provides the possibility, if necessary, of performing sub-analyses to determine the effects of participation in the training program on these Ss from the higher COMPSTAT levels.

2. Establishing the Initial Equivalence of the EXPGROUPS

One set of statistical analyses had to be performed on the COMPSTAT data for the Low competent Ss before the Training Program could begin. It was necessary to establish first that the four EXPGROUPS assigned to the treatment and control conditions did not differ significantly in their pre-training scores on the major variables of analyses. That is, we had to demonstrate the similarity, or initial equivalence of the baseline starting positions of the groups on these measures if we wised to claim later that post-training changes in performance levels were really due to treatment effects. Accordingly, a series of univariate analyses of variance were performed for the four Low groups on all of the measures that were used to evaluate the COMPSTAT results for N = 97. Tables 2 - 38 show the results of these analyses.

The results of these tests of Initial Equivalence can be summarized very briefly. None of the 36 tests showed any EXPGROUP differences except for two STL measures -- the basic ratio measure of Quality corrected for difference in message length, and the log transformed measure of message Efficiency. In both of these latter cases, groups III and IV appear to have higher scores than groups I and II, but these may be due to underlying differences between the groups in message length rather than in message quality. The same qualification can be applied to the only differences encountered among the LL variables. None of these LL analyses produced any EXPGROUP differences, but there were three significant Position differences for the Discriminability Index, the basic, untransformed measure of message Quality, and the regression measure of Quality corrected for differences in message length. However, the direction of differences by Position (B > A) is consistent for all four groups, and since there are no EXPGROUP or Interaction effects, initial equivalence can still be claimed for the four EXPGROUPS. As we have seen in the earlier univariate analyses for N = 97, these Position differences may be accounted for by difference in the amount of propositions generated in one Test-Position as opposed to the other, with the differences in propositional quality being subordinate to this. Therefore, across all of the domains and across all of the measures, with only two exceptions, initial equivalence of pre-training performance appears to be established for all of the EXPGROUPS.

3. The Experimental Predictions

Now that the COMPSTAT groupings have been shown to be significantly different, and the EXPGROUPS performance has been shown to be similar prior to training, we can lay out the experimental predictions that will be tested by the Training Program. There are three Experimental Contrasts of Group Effects. The first contrast tests the effects of Practice, while the other two test the differential effects of Modeling.

Contrast A will compare all of the Treatment groups (I + II + III) against the post-training performance of the Control group (IV) to measure any significant differences between the groups as a result of training. The other two contrasts test the differential effects of peer-modeling within the three EXPGROUPS. Contrast B compares groups I + II vs. III, and Contrast C compares group II vs. III to determine if there is any significant benefit to matching-up children of different needs and abilities with appropriate peer-tutor

If neither of Contrasts B and C proves to be significant, while Contrast A does, this will indicate that participation in any communication training program is better than none at all regardless of the peer models involved. On the other hand, if either or both of Contrasts B and C prove to be significant, the study will have gone a long way towards answering the problematic educational question of "match-up" -- that is, given children with different levels of communicative competence, what is the most efficient and beneficial pairing of language models and learning experiences for children who are less competent in the area of pragmatic-linguistic functioning?

One caveat needs to be entered into the record regarding predictions for the effects of training. The predictions formulated for the original research proposal were based on the characterization of competence predicted by the original tri-partite model of competence. Now that the Evaluation results have further differentiated the language component of that model into two separate functions, it may now be necessary to further differentiate the experimental predictions regarding the effects of training on the separate LL and STL functions. It is entirely possible that training will not benefit the children's functional language all across the board but, instead, will improve one pragmatic function more than the other.

Be that as it may, the most conservative set of predictions would state that the treatment groups would perform at a higher level after training than the control group, and that the differential effects of peer-modeling would follow a predictable H > M > L or III > II > I order within the three EXPGROUPS.

PART II

Training of Communicative Competence

PART II

TRAINING OF COMMUNICATIVE COMPETENCE

II. 1. Overview of the Training Program

The second major phase of this study deals with the attempt to design and implement an educational program for fostering the development of communicative competence in the target schools. As originally conceived, this part of the study was intended to be an exercise in applied psycholinguistics wherein the rationale and results of the Evaluation Phase (Part I) are carried one step further. Can we, in essence, use the psychometric and diagnostic evaluation data to guide and direct an intervention into the pragmatic language development of selected schoolchildren? Can we create classroom language environments with communication tasks which enable children who score low on our tests of communicative competence to practice and develop their speaking and listening abilities? And can we base this practice on dialogue situations where such children are exposed to, and respond to the communicative speech of their peers and adults?

We attempted to answer these questions in ways that were consonant with the philosophy of the model of communicative competence discussion in Part I, and which followed as closely as possible our working definition of competence. The structure of the resulting training program and the curriculum development it necessitated will be discussed below. First, the major features of the program will be discussed as they relate to the experimental goals and hypotheses of the study. Secondly, the overall structure and execution of the training program will be detailed. This will be followed by an exposition of the basic philosophy and content of the program, illustrated via the curriculum of communication games which was especially developed for this study. Finally, a concluding discussion will attempt to deal with some of the educational issues raised by the implementation of this program.

II. 1. i. Major characteristics of the training model

Before describing the how and the what of the training program, it is important to consider the fundamental features of this program in relation to the experimental goals and hypotheses of the study. There are basically three major components in the design of the training model used here. The first concerns the role of practice in fostering increased communicative competence and this is studied via the device of forming experimental and control groups of subjects for comparison purposes. The second component is the role of modeling as an influence on the acquisition of competence and this is studied by varying the composition of the members within the different experimental groupings. Next, there is the important issue of the nature of the language environments in which the practice and the modeling will take place. Each of these issues will be discussed in turn.

II. 1. i. a. Practice

The issue of practice effects and the formation of training (experimental) and non-training (control) groups of subjects addresses directly the first of the three primary experimental contrasts described in Part I of this report; *viz.*, can groups of experimental *Ss* who receive training in oral communication significantly improve their communicative competence when compared, as a whole, with a group of control *Ss* who do not receive such special training?

The experimental purpose, then, is a planned intervention into the oral language experience of certain pre-selected schoolchildren. This intervention has two main foci.

First, the target *Ss* for the study are all those children who have been identified as being "Low Competent" in the pre-training evaluations. These Low *Ss* make up the experimental group of *Ss*.

The other emphasis is on developing the *expressive* language abilities of these Low *Ss*. Even though the training curriculum is intended to integrate as much as possible the complementary processes of speaking and listening and to give equivalent amounts of practice in both oral production and comprehension, it is the expressive language data that will eventually be the primary focus of analysis in the post- training evaluations.

The type of communication practice envisioned for the Low *Ss* is illustrated in Figure 5. All of the experimental groups of Low *Ss* (1, 2, and 3) receive communication training, while the control group of Low *Ss* (group 4) do not participate in the special program but, instead, receive their regular class lessons. All of the experimental groups receive adult modeling in their training sessions, while two of the groups (groups 2 and 3 in Figure 5) also receive exposure to more competent peer models (as will be discussed below).

The conceptualisation of the training program was motivated by the same underlying philosophy and model of communicative competence as inspired the construction of the evaluation battery in Part I. There is, however, one very important distinction to be made here concerning the different aims of the experimental strategies employed in Part I and Part II of the study. The distinction is necessitated by our differing conceptions of the contrasting roles that the three component processes (perceptual, cognitive and linguistic) must play in the Evaluation and Training phases of this study. The intrinsically different goals of these two phases led us to maintain a critical distinction between structure/function and analysis/synthesis of competence.

On the one hand, the analysis or structural description of communicative competence into its component parts is an appropriate prerequisite for the scientific measurement and evaluation of that competence. On the other hand, for the purposes of teaching and learning communicative competence, it seemed to us more appropriate to consider the functional operation of its components in unison, in real communication situations. That is, while it was deemed scientifically proper to study each of these three processes seperately for the purposes of evaluating communicative competence, no such fragmentation of these intricately interrelated psycholinguistic processes was entertained in the training program. Indeed, within the philosophy of the training program, it would be considered a behaviorist fallacy to think that a complex behavior like oral communication could be trained in the same piecemeal, fragmented manner by which it is analysed for purposes of scientific description. Thus the training program, while it recognises and exemplifies the essential role of the perceptual, cognitive and linguistic processes in the act of constructing a verbal message for another, it never engages the children in separated sets of perceptual exercises, cognitive exercises or linguistic exercises. Creating the conditions for the functional employment of these processes in acts of interpersonal communication requires a validation quite different from that used to justify a structural description of the components of competence.

In our training of oral communication, therefore, we take the position that these necessary perceptual, cognitive and linguistic processes operate simultaneously and in parallel in the act of constructing a verbal message about a non-verbal event for a listening other. That is, these processes operate as part of a unified organic whole whose best expression is the speaker's utterance which, in turn, best reflects his or her communicative intent. As such, these processes cannot be meaningfully separated in a program designed to foster and enhance creative and coherent communicative acts on the part of the child-speaker. Rather, the latter should be given opportunities (provided by both the structure and content of the classroom language environment) to activate all of these processes together as the best way of ensuring his or her acquisition of an integrated, balanced competence.

With this wholistic perspective of the learning of communicative competence, therefore, all of the communication tasks and games in the program were constructed to stimulate and activate the child's perceptual, cognitive and linguistic processes in an integrated, unified manner for the purpose of oral, social communication. This aspect of the training model will be further emphasized in the discussion below of the philosophy and content of the communication curriculum.

II. 1. i. b. Modeling

While the issue of "practice" concerns the amount of special training that one group might receive vis-a-vis another, the question of "modeling" is addressed to qualitative variations in the *kind* of practice in oral communication that the experimental *Ss* receive. It is in this respect that the other two primary experimental contrasts in this study are concerned with the differential effects on a subject's performance of being paired within the experimental training groups with peer models of differing levels of communicative competence.

These qualitative variations in competency status can again be illustrated with reference to Figure 5. All of the experimental groups (1-3) received communication practice and adult modeling, but two of these groups (2 and 3) contained pairings or groupings in which Low *Ss* were exposed to higher quality pragmatic language models in the form of Moderate competent or Higher competent peers. The actual breakdown of the sample in terms of competency status (COMPSTAT) and experimental treatment (EXPGROUP) is further illustrated in Figure 4. The core experimental group of 49 Low *Ss* is indicated in row 1 of this table, and it is the data from this core sample that provides the primary analyses and results of this study (Part III of this Report).

As Figs. 4 & 5 indicate, the major difference between experimental groups I, II, and III is in the type and quality of language modeling to which the Low *Ss* are exposed in each group. In order to avoid a "zero modeling" situation with the Low *Ss* in group I, they, as well as the Low *Ss* in groups II and III, received a constant standard amount of interaction with an adult model (who was one of the two teachers trained for this project). However, when it came to child-child interaction, whether in dyads or small groups, the treatment of the three groups varied. The Low *Ss* in group I continued to interact and dialogue with other Low *Ss*, but the Low *Ss* in groups II and III had as their partners Moderate and High competent peers, respectively. In effect, therefore, these differing training formats create "peer-teaching" situations which help us assess both the benefits and the utility of cross-ability matching for regular classroom practice.

We also intended to provide each Low subject with the opportunity to work with as wide a range and variety of communication models as possible, within the constraints of the different experimental treatments as outlined above. So, within each of the three treatment groups the child partners of the Low *Ss* in the child-child situations were allowed to vary within the limitations of membership of each group. Thus in group III, for example, each Low subject had the opportunity to work with every one of the High partners at different times throughout the program.

II. 1. i. c. Language environments

Finally, the communication practice and the language modeling were required to take place in situations that were structured enough and planned enough so that real dialogue would be elicited, that speaking and listening activities would be sustained, and that real learning would take place. In planning the communication training program, we, like the Bullock Report, did not believe that "the child's language can be left to take care of itself" (Bullock Report, 1975, p. 145). As was emphasised in Part I of this Report, we wanted to ensure that the children were talking to some purpose. While we were concerned that their speaking and listening were going on in as naturalistic a context as possible, we were also at pains to ensure that these contexts,or language environments, would make challenging enough communicative demands on the participants so that the program would foster a child's ability to use language as a vehicle for learning.

Accordingly, the program is built around a curriculum of communication games, where the 'game' format allows us to pose problems whose solutions require a verbal exchange of information between the players. Our orientation here, once again, finds an echo in the Bullock Report:

> The important question to ask is whether demands are being made upon their language by the nature of the problem and the process of arriving at a solution to it. Children need to represent to themselves and others what is being learnt.
>
> (Bullock Report, 1975, p. 165)

We sought, therefore in our child-centered, task-oriented games, to create a kind of *dynamic tension* by putting children in situations where one could create and observe the interplay of communicative intentions, where the speakers had communicative *needs* and the listeners imposed communicative *demands*.

In order to create and sustain this kind of pragmatic interplay over a six month training period, we relied on our working definition of communicative competence to help us construct a sufficient variety of dialogue situations that would stimulate a functional variety of language among the experimental *Ss* and their models. Problems were presented in a variety of ways and materials, and cooperative solutions were pursued by the target child and his partners.

The communication training program thus provided the experimental *Ss* with structured and semi-structured, quasi-naturalistic communication situations wherein they could exercise their speaking and listening abilities for social, functional, purposeful dialogue. The atmosphere of these kinds of language environments is non-didactic, and the children learn how to use language simply by using it. The emphasis is on the ongoing *process* of communicating, and not so much on achieving solutions or "right answers" (even though the problems and their solutions provided the *raison d'etre* for most of the games, and were understood and enjoyed as such by the children).

Although we cannot always justify it empirically, it is our impression that the kind of language learning that takes place in these kinds of "live dialogue" encounters is due more to the psychological processes of emulation and assimilation than it is to any kind of "instruction" per se.

More details of the communication curriculum will be provided in Section II.3. below and in Appendix II.A.

II. 2. The structure of the Training Program

After the Evaluation Phase of the study had categorised all of the tested Ss into different levels of competency, the next step was to apportion these Ss to training and non-training groups and get the daily training classes under way. Although it may look simple on paper, dealing with the nuts and bolts of the program implementation proved to be a very complicated, time-consuming procedure. While we sought to implement the ideals and goals of the project as faithfully as possible, we also had to operate within the myriad of constraints imposed by a situation wherein almost one hundred subjects were spread across nine classrooms in three schools in different parts of town. Since the limited budget and resources of this project permitted the use of only two full-time teacher-trainers, some very tight scheduling and coordination of timetables was obviously called for. Since the training program was to get under way right after the Christmas vacation, most of the prior month of December was spent devising daily and weekly training schedules which were compatible with the normal demands of a second grader's very busy school day. On top of this, we had to assign subjects to groups in as balanced and as representative a way as possible to provide a heterogeneous mix of sexes, classrooms and competency status.

The results of these endeavors will be discussed in the following two sections on the formation of training groups and the logistics of program execution.

II. 2. 1. Formation of the Experimental Groups

Using the test results, experimental and control Ss were selected and assigned to training or non-training groups. As reported in Part I of this Report, of the 137 Ss who completed the pretraining evaluations, 50 were designated as Low, 62 as Moderate, and 25 as High competent Ss. Since children who scored low in communicative competence were the targets for training, all 50 Low Ss participated in the program. Within the constraints of school and classroom, they were randomly assigned to three experimental groups and one control group in approximately equal numbers (see Figure 4). Next, to provide peer models for Experimental Group II, 24 Moderate Ss were randomly selected from the pool of 62 such children to be participants in the program. Then a further random selection of 12 of these 24 Moderates was made and these dozen Ss assigned to be the partners of the Low Ss in treatment group II (Low/Mod). The remaining 12 Moderate Ss were designated to be control Ss for the Moderate experimental Ss in group II.

Since there were only 25 Ss identified as High Competent, and all of them needed to be used in the experiment, it is obvious that their participation could not be based on any random selection. However, their assignment to either a High Experimental or a High Control group was done on a quasi-random basis, again within the constraints of school and classroom location since these latter factors predetermined the distribution of such subjects across the school population. In order to avoid having all of the experimental Ss in one school and all the controls in another, a compromise had to be struck, therefore, between randomness and balanced representation. Thus, as described for the Moderate group (above), 12 of the High Ss were assigned to be peer models in treatment group III (Lo/Hi) and 12 were assigned to a control group.

The design of the training program is illustrated in Figure 4 and it is further broken down by school and classroom (or teacher) in Table 1-a. There is a core sample of 49 Low Ss (a Low S in group III (Lo/Hi) dropped out of the program in mid year when he left the school district). The dozen or so Low Ss in each of the three Experimental groupings are matched against one Control group of Low Ss (n=13). Similarly, each of the Mod and Hi Experimental groups is matched against its respective Control group. The total experimental sample comprises 97 Ss. As **Table 1-a** indicates, they are distributed fairly evenly across classes and schools with School A having 34 Ss, School B having 28 Ss, and School C having 35 Ss.

II. 2. ii. Logistics of the Training Program

Execution of the program required a complicated balancing of training variables against subject variables. The training variables included: frequency and duration of sessions, optimal group size and composition, dyadic vs. small group exercises, amounts of adult- vs. peer-modeling, and so on. The subject variables all had to do with the extraordinary busy-ness of a schoolchild's day! In addition to their regular lesson-time in class, children were always moving around the school en route to gym, music or art classes. Sometimes a whole class would vanish for part or all of a day because of a field trip or the showing of a special film. And always there were children being pirated away for special programs -- Title I, ESL, resource room, remedial reading, testing, etc. As a counter-reaction to many of these interruptions of the children's class time, most of the teachers had established "do-not-disturb" periods (usually early in the morning) which they devoted to instructing the whole class in basic skills. ~~Finally~~, as far as the school principals were concerned, the most important variable of all was the availability of physical space to house our training groups. Due to the experimental nature of our research, we required that the trainee subjects be taken out of their classrooms to participate in the communication sessions in a room or area (e.g., half of the library) reserved for that purpose. On certain days in certain schools, this training room might be moved two or three times.

All of the foregoing is a necessary epilogue to highlight the nature and abundance of real-life constraints on scheduling of activities in public schools. Our training program was one more activity that had to be tacked on to the teacher's (and child's) calendar. We would advise anyone who intends to implement or replicate a program such as ours in the schools not to take these constraints lightly.

Details of the daily and weekly operation of the training schedule will be presented next.

II. 2. ii. a. Scheduling of training

The two project assistants who conducted the training in the three schools were experienced male teachers who were especially trained for this program. By very careful budgeting of their time, these two teachers were able to conduct the communication sessions for four days a week, Monday to Thursday, between 8 a.m. and 2 p.m. per day, in each of the three schools. One teacher had primary responsibility for training at School A, and the other teacher for School C (**Table 1-a**), while both teachers shared the training duties at School B. On a typical day, for example, the teacher at School C would conduct training sessions there from 8:30 a.m. until 10:50 a.m. He would then go across town to conduct sessions at School B from 11:15 a.m. until 12:45 p.m., whereafter he would return to School C for the final sessions of the day, from 1:00 p.m. to 2:00 p.m.

Afternoons after school and Fridays were devoted to record-keeping, curriculum development and data management back at the project office.

II. 2. ii. b. Frequency of training

In the initial design of the curriculum each teacher was supposed to work with three or four pairs of children per training session, lasting approximately 40 minutes, on three days a week between Monday and Thursday. Such a design would have allowed each experimental subject to participate three times a week in the program for a total of about two hours per week. However, this design created serious difficulties for the program at the outset, and it had to be modified for a variety of reasons. As it turned out, this initial format was more appropriate to later stages of the program when the children had accumulated a repertoire of communication games and were more able to work independently in small groups, with little or no teacher intervention. The problems we encountered with our initial training format have important, practical implications for the future classroom applicability of this program.

At the beginning, having one teacher work with eight subjects in a 40 minute period created serious classroom management problems which interfered with the quality of the program. The children obviously needed a good deal of instruction and direction as they were being initiated into the program, but the initial format of the program limited the amount of time the teacher could devote to any one child or pair of children. In addition, we discovered that it was often difficult to coordinate sessions as long as 40-45 minutes with the children's other classroom activities.

Accordingly, a revised training format was soon introduced into the program. The children still came three times a week between Monday and Thursday, but the size of the groups was reduced to between one and three pairs per session, and the length of a session was reduced to half-an-hour. However, the number and frequency of sessions per day was increased, with a resulting increase in the amount of teacher-child interaction on a one-to-one basis. Although a child now spent less total weekly time (1 1/2 hours per week as opposed to the original two hours per week) in the program, he was receiving relatively more oral practice time since the interaction and attention he received were more concentrated and more frequent in the smaller group.

For the teacher, the new format meant that, instead of working with eight children during a 40 minute session, he might, for example, work with these same children in three separate 1/2 hour sessions containing groups of 3, 3, and 2 children apiece.

It should be understood, however, that group size or composition was never a rigid, inflexible matter. As the program progressed, as the communication curriculum was developed, and as the subjects became more sophisticated at game playing, extra children might be added to a session whenever the communication task or situation required a larger group of players. This was especially true of small group, face-to-face games like Bingo or Round-Robin storytelling (see Appendix II.A) which might require four or five players.

II. 2. ii. c. Duration of training

The training program covered a five month period from early January to late May, after which the post-training testing took place until mid-June, shortly before the end of the school year.

Of the 19 or so weeks available for training sessions, classes were not in session during two vacation weeks, and one report-card conference week. During the 16 weeks of classtime that remained, the maximum possible number of daily, half-hour sessions that could be devoted to training was 47 per subject for each of the three schools involved. For the target group (see Figure 4) of 36 experimental Low Ss, therefore, the maximum amount of possible training they could receive in toto was 1692 half-hour sessions, or 846 hours of training. This amounts to 23.5 hours of communication training per subject on average that was potentially deliverable.

We, therefore, embarked on the task of achieving this ambitious training goal, and log-books and record charts were kept of the daily training sessions to help us monitor our progress. Due to holidays, snow days, illness and other unforeseen interruptions of a child's training schedule, of the total number of training sessions possible we were able to deliver in actual practice 1416 half-hour sessions, which represents 83.7% of the maximum possible. This is equivalent to 708 hours of training by the two teachers for the 36 experimental Low Ss. The mean amount per child was 19.7 hours, ranging from a minimum of 18.9 hours to a maximum of 20.1 hours on average. For individual children, the least amount of time in training was 17 hours, while 23 hours was the maximum any single child achieved.

II. 2. ii. d. The structure of the experimental group and the role of the teacher

This section will briefly explain the groupings and pairings of experimental Ss in the training sessions and the nature of the teacher-trainer's participation. However, in order to provide a comparison between the subjects' experimental treatments and their regular classroom experiences, it will be instructive to consider for a moment the kinds of classroom environments in which the control Ss remained throughout the study and in which the experimental Ss passed the bulk of their day.

The regular classroom: All of the 9 regular class teachers were interviewed and their classrooms observed on an informal basis. This revealed that the subjects in the study were exposed to a very heterogeneous mix of teaching styles and classroom structures. Classroom organization ranged all the way from traditional, highly structured groupings to more loosely-knit, open systems. Some teachers preferred to work with all of their children in a group for most of the lessons, while others ran a more individualized curriculum. Still other teachers operated with a mixture of both of these approaches, breaking the day up into large group periods and small group or individualized activities.

Attention to speaking and listening varied greatly from class to class. None of the teachers used a formal oral language curriculum per se. For the majority of them, "language arts" was heavily oriented towards phonics, reading, writing and spelling. Opportunities for child talk were quite limited in some classes, yet encouraged in others. We observed a good deal of teacher monologue or limited dialogue where children answered teachers' questions. Lacking a formal or systematic oral language curriculum, one or two of the more experimental teachers relied on improvisation and eclectic sampling of materials to ensure that their pupils received some regular practice in expressive language activities. These included story-telling and tape recording, play-writing and acting, puppetry, etc.

Viewed from the context of the present study, most of the nine classrooms presented two noticeable features: a truly child-centered oral language program was rare, and (as the Bullock report has also noted) there was a lack of sufficient integration of speaking and listening with reading and writing.

The experimental classroom. As indicated earlier, the subjects were taken out of their regular classrooms according to a pre-arranged schedule, and they received their daily training session in a room or area reserved for that purpose. Depending on the size of the group or the kind of communication exercises designated for that day, a low experimental subject might work in groups ranging in size from a pair (a child-child or child-adult dyad) up to four or five members.

The choice of partners for the experimental subjects was predetermined by the treatment groupings (Figure 5). However, rotation and variation of partners within groups was systematically controlled on a daily and weekly basis. The log-book records of the daily training sessions for the three schools show that the planned rotation of experimental Ss across peer or adult models was successfully achieved. Even with the constraints of scheduling, every one of the 36 Low Ss in treatment groups I, II and III was able to rotate across the full complement of partners in his or her modelling group (Low or Mod or High) in a regular, cyclical fashion. For example, subject #076, who was in treatment group II with Moderate Competent peer models as partners, had the following rotation of partners over a typical five-day period: first day -- S#080 (male peer), second day -- S#043 (male peer), third day -- S#067 (female peer), fourth day -- S#085 (female peer), and fifth day -- adult male teacher.

Over the five months of training, our goal was to provide an equal balance of adult modelling (constant) and peer modelling (Lo or Mod or High) for the 36 Low Ss in the three treatment groups. Absenteeism by subjects or their partners (for whatever reason) naturally interfered with the training schedule, but despite this, the records show that our goal was reasonably well achieved. For all Ss, across all training sessions the proportion of adult -- as opposed to peer -- modelling was approximately 54% (adult) vs. 46% (peer). There is some overlap, however, between these percentages since they do not distinguish between the times when either kind of modelling occurred "solo" (that is, when an adult or a peer was the sole partner of a Low S in a dyad situation) and when it co-occurred in group communication games. An example of the latter situation would be in the game of "Bingo" (Appendix IIA) where the adult teacher-trainer might join the peer models to play a five-hand round of the game.

We also think that another reason for the slightly higher proportion of adult modelling was the change in training format (Sec. II.2.ii.b., above) that was made early in the program. When the sizes of groups were reduced, a larger proportion of odd-numbered groups was created (e.g., three or five). In cases like these, two of the children would form a dyad for a communication game (e.g., Object Assembly) while the teacher would participate as an adult model in a dyad with the remaining child in another game (e.g., Movie Director). In larger group games, the teacher might participate with the whole group as a model (as was the case in "Bingo", above), or he might let four children (say, two Lows and two Highs) play by themselves as a group or in two dyads, while he conducted a one-to-one session in a dyad with the remaining child.

The role of the teacher in all of these instances is always non-didactic. He or she is not a dispenser of information to passive listeners. Given the child-task orientation of the curriculum where the "child talk" is the focus of attention, the teacher is required to play two roles -- first as a language model, and second as a manager of resources. In the first instance the teacher tries to act as a model of higher quality expressive language exposing the child to a richer, more varied and more complex vocabulary and syntax than he might normally hear or use himself in comparable situations. But this modelling of language structure is always in the service of function, and the pragmatic, communicative setting of the teacher-child exchanges in this study ensured that the adult is always modelling how language is used. The structure and dynamics of the communication tasks then afford the child prompt, timely and frequent opportunities to practice the new forms and expressions he hears and to incorporate them into his own repertoire.

Given the variety of tasks, groupings and partners in an oral language curriculum such as ours, the teacher's other role is to manage the resources of the program, orchestrating and conducting the communication activities. The teacher "directs traffic" as it were, and helps learning to occur in ways that avoid didactic instructing on his or her part. For example, impasses in the flow of communication between children are resolved by the teacher's providing judicious questioning or subtle and indirect suggestions or hints which are all aimed at fostering and exercising the child's capacity for self-reflection. This ensures that the solution to the particular communication problem is more likely to be evolved by the children themselves and to be carried over as a learned experience to other communication games and activities.

II. 2. ii. e. Curriculum development and the nature of the training activities.

One of the aims of the original proposal was the development of a set of recommendations regarding the major ingredients for an oral language curriculum that is aimed at promoting the acquisition of communicative competence. Accordingly, much of the work of curriculum development was experimental in nature, requiring a good deal of brainstorming and trial and error evaluation during the training phase of the study.

As originally proposed, a set of core exercises or "communication games" was developed to provide a basic curriculum. Variants of each of these core exercises were also designed, with varying formats and levels of difficulty, so that a wide variety of games and exercises, utilizing a wide range of materials and tasks, became available to provide a substantial curriculum repertoire that the teachers could dip into to meet the needs of the daily training sessions. Within the guidelines of the study, our approach to curriculum development was very eclectic. When necessary, the project would adapt commercially available materials for its own purposes. But the major part of the communication training program was designed from scratch, and a great deal of time and creative effort was devoted (outside of school hours) to the development of appropriate training materials whose structure and purpose were derived from and consonant with the rationale and model of communicative competence presented in Part I of this report.

Much of our new material was used on a trial and error basis. Some of it was either discarded if inappropriate or unworkable, or it was subsequently revised and modified in order to become a viable component in the overall training program. Appendix IIA presents a representative cross section of the games that worked.

The next section on the philosophy of the training program will present detailed recommendations for what we came to believe are the necessary ingredients of a truly oral language training program. Before discussing the philosophy of the program, however, it is necessary to consider briefly the communication training activities that reflected that philosophy in the day-to-day operation of the program.

As was emphasized earlier, our activities were based on an approach to language intervention that is essentially "non-didactic". Instead of direct teaching by the teacher, the latter has to create and manipulate the language environment of the classroom in such a way as to allow the child to learn how to use language by using it. That is, instead of the teacher didactically trying to indoctrinate the child with information, the child learns the necessary information and communication skills via verbal interaction with others. This interaction is sustained by our "games" approach to language learning, an educational strategy whose virtues have recently been re-emphasized (Karnes, 1978). The placing of talking and listening tasks in a game context provides a structure and a purpose which the children find highly motivating. This motivation was further enhanced by designing games which created "conditional" or "contingent" communication situations (Chandler, 1975). That is, in order to create the communicative necessity which was the motivating force behind the dialogue in our games, it was necessary to manipulate the balance of information possessed or needed by any or all of the players in a communication game. Thus, one player may be given 100% of the details necessary to conduct a game while the other(s) has zero. Or the ratio on either side may be 50%/50% or 70%/30%, etc... In any case, what effectively happens is that one creates a conditional communication situation wherein the nature, progress and success of the game is contingent upon a verbal exchange of knowledge and information.

All of the communication tasks in the program required, to varying degrees, the necessary exercise of the child's perceptual, cognitive and verbal faculties. These were exercised in the service of various pragmatic language functions, primarily logical language and spatio-temporal language usage (as discussed in Parts I and III of this Report). Depending on the game, the players were required to engage in description, explanation, instruction, direction-giving, cause-effect reasoning, if-then deductions, interpretation, judging, and so on. The training activities also required the experimental Ss to deal with a variety of materials -- toys, everyday objects, abstract forms, playing cards, photos, storybook picutres, etc. The subjects also had to learn to deal with different levels of representation -- "live" models, three-dimensional figures, two-dimensional photos, caricatures, line drawings, etc. The materials further ranged from concrete to abstract levels of complexity.

The social structure of the games -- dyads and small groups -- has already been described, and the details of the games presented in Appendix II.A (picture descriptions, map reading, object manipulation, puzzle solving, etc.) provide sufficient illustration of the variety of communicative demands placed upon the children.

In sum, the nature, structure, content and goals of our speaking and listening activities had a common focus, viz., to require the child to work in situations where he experiences at first hand the need to communicate, and where communicative demands are placed on him by others in order to challenge his expressive abilities. To this end, the training tasks presented the child with communication problems which had clearly defined goals and which made specific demands on the child to listen to and talk to others. Thus, as was emphasized above, we were concerned not simply with openended discussion, but with talking to some purpose.

II. 3. Philosophy and content of the training program[1]

The content of the training curriculum is presented in Appendix II.A. The principles and recommendations which should guide the implementation of such a program are presented next.

II. 3. 1. Some opinions and assumptions

a. Education-as-right-answer versus Education-as-process?

This research project has operated on the assumption that the full and proper development of the child's oral language is a necessary pre-requisite for the optimal acquisition of later forms of literacy such as reading and writing. Because of this belief in a solid oral base for all forms of language learning, we also contend that the core of a language arts communication program should be active use of language. As a consequence, therefore, our own training program is very child-centered and its emphasis is on children talking. The concern is with the process of verbal message construction and with the pragmatics of language development -- that is, how the school child uses his pre-acquired linguistic repertoire for effective interpersonal communication.

Unfortunately, too much of what passes for "oral language practice" in the classroom is really just Monologue (the teacher's) masquerading as Dialogue. We all know that teachers are good talkers, but are they also good listeners? Children can learn how to use language only by being allowed to use it for a variety of social, communicative purposes, both inside and outside of the classroom. The educational problem then becomes one of how to create the learning situations in schools that will make demands on the child to communicate and foster his ability for oral expression and comprehension.

To foster the child's language usage skills, teachers must involve the child more deeply and more fully in oral dialogue in the classroom. To do this, many teachers will have to revise traditional conceptions both about what the child is capable of as a language-learner and about the nature and structure of meaningful oral language experience in the classroom. A positive step in this direction would be taken if teachers would operate on the basis of two simple assumptions. First, young children enter school with a pre-acquired basic linguistic repertoire -- that is, they already know a good deal about their language. Teachers can and must take advantage of this knowledge and capitalize on it. Secondly, young children also enter school, however, with wide individual differences in their "communicative competence", or their ability to use what they know for pragmatic, purposeful communication. To counter these differences, the teacher must make the language environment of the classroom structured enough, stimulating enough and talkative enough to develop each child's communicative competence.

Our aim here is to show teachers how they can create a language environment in the classroom wherein the child is stimulated to communicate in a variety of ways, for a variety of purposes, in a variety of situations, to a variety of listeners. In talking of such language usage, we are talking, therefore, about the child's production of connected discourse, and not of his emission of pre-determined responses to adult questions. Yet, for many teachers, "learning" means learning "right answers" (often one-word replies) and learning a language in particular means learning a bunch of "verbal responses" which can be changed or "corrected".

---

[1] This section is adapted from an article published in French in the Canadian education review, Quebec Francais, No. 25, March, 1977, under the title "Parler en classe: une maniere d'aborder la communication orale dans les classes elementaires, de facon non-didactique."

But this approach to learning and teaching (education-as-right-answer) ignores the fact that behind the child's answers there lies a whole complex of psychological and linguistic processes which govern and direct the child's use and understanding of the speech of his community.

If we can change the teachers' focus from overt "response" to underlying "process" we may be able to kill two birds with one stone. For example, on the one hand, such a change in teaching attitude may create a more genuinely "child-centered" pedagogy by making teachers more sensitive to developmental aspects of language learning. Thus, their teaching may become more profitably attuned to individual differences in rates and levels of learning, enabling them to work with the strengths as well as the weaknesses in the language capacities of their students. On the other hand, a concern with "education-as-process" may also lead us out of the morass of "corrections", "tests", "scores", report-card grades", etc...by turning our attention more towards long-term evaluation and the passage of the child through various "stages" of development or mastery or competence in the use and understanding of his language. Hopefully, therefore, by trying to understand the psychological and linguistic processes which influence the child's verbal "responses", the teacher will be able to develop an oral language curriculum that is more responsive to the needs and capacities of the child and less dictated by the needs and politics of the educational "system" in which she has to operate.

b. Oral communication as a cooperative enterprise

If the teacher can now come to acknowledge the linguistic strengths which the child brings with him to the classroom, our hope is that she can also come to realise that, in so far as oral communication is concerned, the child should now play a more active role in his own learning. This does not mean that the teacher's role should now be a more passive one, but rather, a different one. For a dialogue needs two active partners, and, as the Bullock Report sanely points out (Bullock, p. 145), the child's language cannot be left to take care of itself. In order that there be a true, balanced, productive dialogue in the classroom, at least two conditions must exist. One partner in the exchange (the child) should be allowed (indeed be solicited) to make an oral contribution which will be accepted and respected and responded to by the teacher. But to complement the child's role, the teacher, as Bullock suggests, must play a role of "planned intervention". (Bullock, p. 145)

Yet in order to be planned, this intervention must also be informed, and it is here that new research, like the present study, can help educate the teacher about the principles and processes of oral communication in children. Language education requires an understanding of not only the psycholinguistics but also of the pedagogy of the problem. As our study has attempted to translate theory and research into classroom practice, there has evolved a set of pedagogical principles which we can offer to the teacher who wishes to plan a language intervention program. These principles, if understood and applied correctly, will enable the teacher to create the kind of language environment in the classroom that provides an opportunity for dialogue and for the development of oral expression as a cooperative enterprise between the teacher, the child, and his peers.

II. 3. ii. The basic principles

We will now outline a set of ten or so basic principles which will provide the fundamental rationale for creating a purposeful "dialogue" environment in the classroom. Armed with this finite set of principles, the teacher will be able to create an infinite variety of communication games and exercises which she can use to carry out the goals of the oral language program.

We are confident that these psycholinguistic and pedagogical principles will, when implemented faithfully and honestly, assure the creation and maintenance of a classroom language environment in which dialogue can flourish. The role played by these principles in actual practice is illustrated in the communication games which were developed for the experimental training program and which are presented in Appendix II.A.

1. "Organic"

The first principle requires that the act of verbal communication be treated as an organic entity. Although communication behavior may be artificially segmented into components (e.g., perceptual, cognitive, verbal) for the sake of analysis and measurement, we do not commit the error, as other language development programs have done, of perpetuating this artificial fragmentation of the communication process in a training program. (We have previously alluded to this error, in section II.1.i.a., as the "behaviorist fallacy".) In our program, for example, we do not give separate training exercises in perceptual, or cognitive or verbal skills. Rather, we place the child in contingent communication situations where, in the act of constructing a verbal message for an other, all three of these skill processes need to be utilized simultaneously and naturally and as part of a whole, i.e., organically.

For instance, in all of the tasks illustrated in Appendix II.A., one can identify a number of interlocking processes which govern the speaker's overall response to the task. For instance, the child has to make a perceptual analysis of the situation to be talked about; furthermore, he has to take account, cognitively, of the perspective of the listener; and then he needs to encode these observations into a verbal message form that the other can understand. But these analyses are not separate, sequential operations; they are, rather, simultaneous, parallel processes.

2. "Human"

Language is a uniquely human form of social interaction, and it is in such a context that its use should always be demonstrated to young children who are trying to master and practice the rich variety of human communicative functions that language has to offer. Like the Bullock Report, we wish to emphasize the "inseparability of language and the human situation". (Bullock, 1975, p. 528)

In attempting to achieve such an orientation in the classroom, the type of communication that children engaged in throughout our training program was always human, with live partners, usually in "face-to-face" situations. Unlike other language development programs, they were not umbilicised to cassettes and tape-recorders, listening alone to disembodied voices in so-called "decoding" exercises.

3. "Modeling"

Closely tied to the above principle is the use of "modeling" and peer-teaching in an oral communication program. That is, other children, as well as adults, can serve as language models for a child who needs help with his language usage skills.

We have previously discussed (above) this use of modeling rather than didactic techniques to improve the linguistic and communicative quality of the child's verbal messages. This approach is based on the belief that the acquisition of language structures and language functions by the child both have one thing in common; they are both facilitated by exposure to competent models who know and use the language well. The educational role of models is recognised in various ways in our training tasks to ensure that a child can learn the functions of language by hearing other people using language in a natural context.

For example, in our communication games, one of the players can easily be the teacher when a child needs to work with an adult model who can help sensitise him, say, to the use of a more precise vocabulary. Or again, in order to free her to work with other children with special needs, the teacher can create groups of children of mixed abilities who will learn from each other while playing the communication games. Thus, the teacher can change the organisation and management of her class by using her own pupils as teaching resources.

4. "Practice"

The next principle operates in conjunction with "modeling". It is practice -- the child essentially learns by "doing", by getting constant practice in constructing verbal descriptions, explanations, instructions, etc., for specific purposes in achieving cooperative solutions to communication problems with other partners.

There are at least three distinct practice effects and benefits operating in our curriculum. In the first instance, just alluded to, the structure and dynamics of the training games give the child plenty of practice in constructing a variety of oral message forms for others. If these messages are successful then the value of the practice is practice itself. On the other hand, the effects of feedback (in the form of questions, wrong responses, blocked action, etc.) may force the child, through an act of self-reflection, to spontaneously modify and transform his original messages. The continuing dialogue exchange with his partners allows him to practice immediately these modified message forms.

Thirdly, if modelling exposes the child to new utterance forms he is unlikely to produce spontaneously himself, then the nature of the training program allows him prompt and timely opportunities to practice using these new forms in a social, communicative context.

This principle of practice, and its companion, modelling, have wide-ranging, general utility and value in any oral language curriculum. However, it should be remembered that they also have a very specific experimental purpose in the present study. The question of the benefits of practice is addressed by the first of the experimental contrasts between experimental and control groups. The value of modelling is then evaluated in the other two contrasts within the experimental groups. The results of these contrasts are presented in Part III of this report.

5. "Feedback"

In conjunction with the latter two principles, "feedback" must be an essential element in an oral communication curriculum in order to ensure that learning is taking place. A continuous feature in our communication training sessions is the use of feedback and "cognitive discrepancy" to make the child aware of any mismatches between his messages and the effects they are intended to create.

He will thus have the opportunity to modify or rephrase his message while the exchange of information is still going on. The form and the purpose of the feedback will vary from game to game. In Movie Director, for instance, the feedback is immediate, while in Object Assembly it is "delayed", in the sense that the speaker must first finish his message before coming around the screen to see the results of his instructions.

6. Integrated "Talking" and "Listening"

In the kind of program we are recommending here, it is absolutely essential that the teacher avoid false divisions between "talking" and "listening". In all of the communication tasks in this program, each child plays a dual role -- the "decoder" (the one who listens) as well as the "encoder" (the one who speaks). Thus, each learns not only how to speak effectively, but also how to interpret and handle a variety of message types received from others.

In a game like Movie Director for example, the speaker not only has to give directions to the "actor", but he must also attend to and interpret the questions of the latter in order to modify and/or improve the quality of his own messages.

7. "Match-Up"

A continual problem in a wide-ranging oral language program is flexibility -- how to have the variety and depth of activities so that the teacher can find the best "match-up" between the levels of a curriculum and the needs of a particular child. The curriculum-child "match-up" problem in our program is tackled by using tasks, games and materials which are especially designed with a hierarchical order of difficulty built into them so that the level of a communication problem can be matched to an individual child's needs and communicative capacities.

Each of the games illustrated here, for example, is flexible enough to be used at several different grade levels. While the format of Movie Director or Object Assembly may stay the same, the teacher is free to vary the complexity of the photos and pictures, or the familiarity of the objects, in order to make the communication problem easier or harder. Games that can be played at different levels of difficulty allow the teacher to move a child backwards or forwards in a curriculum to a level best suited to his or her current needs and capacities.

8. "Acceptance"

Allied with the principle of feedback, and perhaps central to the whole philosophy of a child-centered oral language program, is our complete acceptance of whatever the child says or offers as a message. There is no didactic correction or criticism of a child's communications, in the kinds of games illustrated here. When the "organic", "modeling" and "feedback" principles are operating correctly in an oral curriculum, this eliminates the necessity for the teacher to intervene in a traditionally didactic manner by way of "giving" the child the "right" answer.

Our advice is to try, instead, to recognize that the child is psychologically equipped to learn language in a functional context, and try to accept anything he offers as a verbal message form as a point of departure.

This type of philosophy created a very positive atmosphere throughout our training program and made fluent talkers out of many children who were previously inhibited. Any "teaching" that was done in the program was done through the use of carefully designed "cognitive discrepancy" situations of the sort referred to above, where the effect of feedback is to inform the child about the comprehensibility of his or her messages.

When the teacher must intervene in a game to control the progress of the dialogue, such "teaching" is effected through the subtle use of questions and hints for self-reflection to the child by the teacher, as well as through the continuous verbal feedback and dialogue that the encoder has with his partners.

9. "Principles, Processes, Ideas"

In designing the content and structure of the training program our constant emphasis has been on the principles and processes of communicative competence and not on the availability of appropriate materials. That is, our pedagogical bias is towards ideas, not technology. Any teacher who knows (i.e., understands) which communication skill it is that she wants her pupil to learn can devise her own materials.

What our program tries to offer to the teacher is a "do-it-yourself" packet of principles which she can use to generate any number of communication games and activities. In this way, we hope that the teacher can be freed from continual dependence on commercial kits and packaged programs, which so often lack a theoretical or empirical justification for their alleged educational value. If teachers are to understand the full meaning and implication of a planned intervention into oral language learning, then a distinction must be made between capricious reaction to available materials and the informed, creative selection of technical material in accord with a pre-conceived teaching purpose. That is, rather than saying "here are some attractive, new materials -- what can we use them for?", the teacher should be saying, "here is the kind of language environment I am trying to create in my classroom, so which of these available materials will be appropriate for my purposes?"

The importance of such a pedagogy based on a priori principles can best be illustrated by quoting a popular saying:

> "Give me a fish and I can eat for a day;
> teach me how to fish, and I can eat for
> a lifetime."

10. "Child-Task Orientation"

One often-cited goal of elementary education is to help the child become an independent agent of his own learning. One means of achieving this goal is to create circumstances that require the child to play an active role in problem-solving situations. To effectuate this in the realm of oral communication, our game-based program adopts a "child-task" orientation, rather than a "child-teacher" orientation. By maintaining a pragmatic, functional emphasis in our design of communication tasks and in the kinds of communicative demands we make on the child, we hope to make him competent (in a communication sense) in dealing with a wide range of situations, materials, problems, and partners. One of the highlights of our small group communication games has been the way in which conditional or contingent communication situations have been used to make the children engage in cooperative solutions, with their partners, to a common problem.

When one follows such a "child-task" orientation in conjunction with cooperative problem-solving, not only do the children find this highly motivating, but they also are forced to listen to their partners in order to complete the task (this is especially true in games like Round Robin Storytelling). Thus, two components of the learning process -- attention and concentration -- are addressed at the same time with such an orientation.

So much for the ten commandments. There remain only two or three points, or sub-principles, to elaborate under the headings of "Dynamic Structure", Manifestation" and "Heuristic".

To the teacher who wants to try to apply the principles outlined above in her class, some advice should be given regarding the delicate balance she will have to maintain in carrying out these activities. The format of the communication games must be structured enough to provide a support for the child, along with a sense of well-defined purpose. Yet, at the same time, these situations must be open enough and loose enough to allow the child to express himself freely and spontaneously therein. Being able to manage such a judicious combination of organization and flexibility is what we characterize as the achievement of dynamic structure.

To the teacher who really wants to listen while her pupils speak, we can assure her (from our own experience and that of other teachers) that the experience will provide two rewards. Because, when a child is required to communicate in a variety of problem situations, the teacher will not only hear what the child does know about language use, but also what he does not know (or knows imperfectly). It is this kind of manifestation of knowledge and ignorance that is a direct result of a program based on talking. It allows the teacher to discern both where she can rely on the child's strengths, and where she must devise special lessons (e.g., specialized vocabulary) for his weaknesses.

Finally, to the elementary teachers whose concerns go beyond strictly oral language programs, we would like to pose one question: Can you conceive of the above set of ten principles as representing a heuristic, or pedagogical paradigm, or general model for a new approach to other areas of the curriculum like reading, or writing, or math, etc.?

In conclusion, we wish to echo the Bullock Report by asserting that the language the child learns in school must equip him for all of his life and not just for the classroom. The scope and relevance of the oral language experiences the child encounters in school must prepare him for this wider dialogue. It is in this context that we have been trying to argue, above, for the creation of the kind of cooperative enterprise, via dialogue, between child, teacher and peers, which would provide a starting point for just this kind of life-relevant language education. For what the child brings to the classroom dialogue must be heard, not only because his potential must be realised, but also because his contribution needs to be transformed by the teacher so that the child, in turn, can transform his own life.

II. 4. Discussion

i. The effectiveness of the program

Having discussed at length the workings of the communication training program, we now conclude with a discussion of the effectiveness of the program. This should be done not only in terms of the goals of the original proposal, but also in terms of the new and unforeseen questions and issues that were raised in the course of conducting the program, together with our attempts at dealing with these.

The original training goals of this study were fully realised: training groups were formed in accord with a prior, systematic evaluation of the subjects' levels of communicative competence; the training program was based on an oral language curriculum which was built around a set of core exercises and developed in accordance with a theoretical model of communicative competence; different experimental groupings of Low competence Ss were used to evaluate the roles of modeling and practice in fostering the development of expressive language abilities; and finally, our experiences with developing and implementing the program enabled us to make some recommendations about important ingredients for an oral language curriculum.

The empirical evidence for the specific effectiveness of the program will be presented in Part III of this report where the post-training assessment results will be discussed. What remains to be discussed here is another more general kind of program effectiveness which was not measured by the battery of perceptual, cognitive and linguistic tests. That is, the effectiveness of the study in raising and dealing with questions and issues that have important implications for other planned interventions into oral language learning in the future. These issues include: a) the choice of philosophy and content of a training model and the companion choice of the activities and exercises that will enhance oral language learning; b) the psycholinguistics of an oral language curriculum -- what is the nature, range and variety of language to be elicited? c) what kinds of learning go on in a program such as ours and how do these learning effects generalize within and beyond the program? d) how should the issue of classroom management be tackled in practical application of this program in schools? e) the need to monitor and evaluate an oral language training program in the classroom: short- vs.. long-term goals.

For some of these issues, parts of our program have made direct and original contributions to understanding. To some of the other questions, our experiences have produced only partial or indirect answers and the need for further research and experimentation is indicated. Let us deal with each of these five issues in turn.

a. Philosophy, structure and content of the training model.

The choice of both the instructional model and the curriculum content are crucial to the essential character and purpose of a language education program. In combining a psycholinguistic and pedagogical rationale with a practical "games" approach to curriculum development we feel that the present program has made an original contribution to the teaching of speaking and listening. We will now take up each of these two areas respectively.

The philosophical approach which we offered in section II.3. above originated with a conceptualization of communicative competence and evolved into a comprehensive pedagogical strategy as a result of our experiences with program implementation. The feedback from practical application of our ideas in the classroom caused us to rework and refine this philosophy to the point where it is now possible to claim that one can see all ten of our basic principles at work in every one of the communication games that were developed.

Each educator or curriculum developer will choose an approach to language education that best reflects his or her theoretical preferences and priorities.[1] The educational philosophy offered here is biased towards cognitive-developmental psychology, especially as this has recently been applied to educational problems (Kohlberg, 1972, Furth and Wachs, 1975, Selman, 1980). It is concerned with the psychological and linguistic processes underlying the schoolchild's practical, purposeful use of language. The resulting program permits observation of the child's pragmatic language in contexts that are structured enough to elucidate the workings of these underlying processes through active problem-solving, and social enough to enhance the development of these processes through communicative interaction. Because of the concerns with the social, interactional contexts of language use, we have, whenever possible, drawn upon recent research in the pragmatics of language acquisition (e.g., Bates, 1976, Halliday, 1973, 1975) to provide support for the psycholinguistics of our model. Unfortunately, as we have argued before (McCaffrey, 1977), the bulk of this new research on language functions has been preoccupied, like previous research on syntactic development, with the infant and pre-school child. Much less work has been done on the development of speech acts in the older, school-age child, and it is hoped that the present study will make a contribution to this neglected area.

---

[1] Karnes (1978), for example, chose for her pre-school intervention program a communication model from the Illinois Test of Psycholinguistic Abilities and a cognitive model from Guilford's Structure of Intellect.

Our experiences in reviewing the language arts and curriculum development literature suggest that there is likewise a neglect of oral language programs in the field of elementary education. There are very, very few systematic and comprehensive programs dealing with speaking and listening in the classroom, and of the few that might claim to do so, hardly any of these have a theoretical or empirical foundation. The Bullock Report (1975) argued for a balanced, integrated language curriculum in schools, with equal time devoted to reading, writing, talking and listening. Unfortunately the two oral language functions appear to be very much the weak sisters when it comes to treatment of these four basic skills ("les quatres savoirs" - McCaffrey, 1977) in elementary schools. It would be hoped, therefore, that speaking and listening might receive more attention from the recent "minimal competency" movement in the U.S. However, in a recent review of state activity in minimal competency testing (Pipho, 1979), only 3 out of 38 states reviewed make specific mention of speaking and listening as skill areas to be assessed. Even where good intentions exist regarding oral language assessment and teaching, the legislative prescriptions far outstrip the available educational expertise in a field such as this where, compared to reading and writing, much of the necessary theory and research remains to be developed. As a consequence, when trying to state policy goals in the area of speaking and listening, the minimum competency movement has been forced, for lack of available knowledge, to couch these goals in terms that are so vague and general as to make them unworkable, or so overly specific as to be pedantic and trite.

Given the present state of the art, therefore, in oral language education, it is hoped that the teaching philosophy and program rationale that were developed in the present study will help educators ask more meaningful and more sophisticated questions about evaluating and fostering speaking and listening competencies in schools.

The second issue that needs to be considered here concerns the choice of activities to enhance oral language learning in the classroom. Our experience in developing and working with a curriculum based on language "games" has shown that such an approach has much utility and benefit. The content of the curriculum will of course be determined by the type of language functions the teacher wishes to focus on. But the structure and dynamics of the instructional format are also important, and it is here that games prove to be a powerful vehicle for learning. Games by their very nature have a built-in metric of difficulty so that a variety of them provides the teacher with a hierarchy of tasks to fit to the different levels of learning styles, needs and capacities of the students.

Our communication games were highly motivating and enjoyable for both the children and the teachers. In addition to their motivational value the games also had a great deal of instructional power. For example, in card games like Bingo or Go Fish or Who Am I? the picture content of the cards might introduce the players to topics in botany, zoology, marine biology, etc. Other social, cognitive and linguistic benefits of game-playing were also readily observable. The children naturally developed their game-playing ability over the course of the program and as their skill developed, so did their confidence in interacting with others. As a consequence they were able to reap the social benefits of cooperative play. From a cognitive standpoint, the games also facilitated a good deal of rule-learning, and the communication tasks constantly challenged the child to process increasing numbers of variables. Indeed the hierarchical nature of the curriculum made this kind of cognitive learning a necessary perequisite for graduating to more complex levels of a game.

A variety of language functions were also acquired or exercised in the course of game playing. Question-asking, for instance, was an important skill that was required and encouraged by many of the tasks. The give-and-take of the dialogue engendered by the game format led the children to think out loud, orally formulating questions, hypotheses and deductions necessary to solve the communication problems. In other games, like Bingo or Who Am I?, where the use of oblique reference was required, the descriptive use of metaphor, analogy and simile were also encouraged.

Games as an instructional strategy have obvious advantages (Lee, 1965), and their virtues in this respect have also been recently extolled in independent research by Karnes and Lee (1978). The game approach is particularly well suited to oral language learning but its use should probably complement rather than replace other instructional strategies available to the teacher. Even communication games at some point have to lead into and integrate with the rest of the language arts curriculum. We have found, for example, that our sentence-building games can only go so far in providing practice in grammatical construction and comprehension. Eventually they have to lead into reading and writing proper.

b. The psycholinguistic content of an oral language curriculum

What is the nature, range and variety of language functions that a teacher should concentrate on in her speaking and listening lessons? We cannot give a highly detailed and specific answer to this question because of a lack of sufficient research on pragmatic language development in the older school-age child. We can, however, urge the teacher to be guided by at least two basic criteria in her choice of language activities. She should first of all ensure that the student is exposed to and practices a representative variety of language usage in order to equip him to deal effectively with the variety of communicative demands that will be made on him both outside and inside of school. Secondly, the teacher should incorporate into the program language practices which build upon and extend the student's prior language learning. The results of the present study show the linguistic effects of following such a course of action.

The original intention in the training program was to provide the subjects with communicative tasks that would elicit a representative cross-section of language uses -- describing, explaining, instructing, recounting, etc. The analyses of the data revealed that such specific language behaviors could be subsumed under two major categories of language function -- Logical Language and Spatio-Temporal Language use.

The former incorporates the use of language for description, classification and categorization of static objects and events, while the latter is used to describe the sequence and transformation of events and actions over time and space. These two language functions are quite fundamental, and it is not unremarkable that their antecedents can be observed in the early acquisition of stative and active verb forms by the infant and young child (Bruner, 1975, Dore, 1975).

A teacher, therefore, might well choose the Logical and Spatio-Temporal Language functions as the nucleus around which to build an elementary, oral language program. There are, however, other options available in the psycholinguistics literature. Halliday (1973, 1975), for example, has theorized about the issue of "pragmatic universals", and has made a sociolinguistic analysis of seven basic and very common language functions acquired by children. Halliday's taxonomy of functions (and their accompanying utterance forms) would suggest a variety of language uses that could be practiced in the classroom. An other, more psychological categorization of language usage is also offered by Tough(1977) to provide the teacher with a framework for identifying uses of language to foster. Tough sees three main categories of usage — Reporting, Projecting and Logical Reasoning — elaborated in further subcategories which provide examples of language practices which can be stimulated and fostered through story-telling and interpretation activities. Tough has developed a pre-school curriculum which incorporates these three major language functions. It is a good example of an approach which exercises a broad, general and representative variety of language usages yet still takes as its starting point the language functions (however rudimentary) that are part of the child's repertoire on entering school.

Finally, it should be noted that the linguistic content of a program cannot be determined independently of other psychological factors. This report argued earlier that the acquisition of communicative competence is not a purely linguistic process. Similarly, the linguistic content of a training program may be enhanced or limited by the type of task that is built into the program. For example, the development of communication games requires a good deal of *task analysis* and an identification of the *cognitive demands* of the tasks themselves (e.g., spatial, temporal or categorical reasoning). The pragmatic-linguistic issue then becomes one of how these demands are dealt with or expressed in the language available to the child.

c. Learning effects and generalization

The effectiveness of this program in promoting language learning among the experimental *Ss* is an empirical issue that is discussed in Part III of this report. Evidence will be presented there to show that significant gains were made on a number of verbal measures. In a study of this sort the question is usually also raised about whether the communication training effects generalize beyond the program to areas of functioning which were not the object of specific training. Part III will present data to show that participation in the program had positive consequences for other areas of cognitive functioning and reading performance. We had also wanted to measure generalization effects with our set of secondary linguistic measures, but, as was explained in Part I, these measures proved to be unreliable and unsatisfactory and were discarded.

Generalization as a psychometric issue is essentially dealt with in Part III. But the training program also was effective in raising important questions about the definition and identification of learning effects that were not directly measured or measurable. It is these more intangible kinds of generalizations that we wish to deal with here.

First, at one level, it could be argued that the type of learning that is measured by the post-training tests is all generalized learning. The test battery consists of separate perceptual, cognitive and linguistic tests while the training curriculum is 100% oral communication, not broken down into sub-routines. Given the philosophy and content of the program there is no way in which the training can be conceived of as "coaching" the Ss to take the tests. Although the screen paradigm is common to both testing and training procedures, none of the verbal, cognitive or perceptual test materials ever appear in the training program. It has been reiterated frequently in this report that we regard communicative competence as an integrative process and that the training program ought to be attuned to and reflect this kind of process. Theoretically, at least, changes in competence as a result of training ought, therefore, to mean changes at the level of the psychological and linguistic processes which underlie the child's specific verbal behavior in specific communication situations. If this analysis is correct, then "new learning" means a change in the status of these underlying processes which determine competence. Furthermore, when he is faced with the post-training test items (which have some similarities with the training tasks but are still relatively unfamiliar), the answers he produces will themselves be generalized responses from his new learning, assuming that such has taken place.

Conceived of in this way, training effects transferred outside of the program provide a special case of generalization. Yet there are other equally important instances of generalization of learning within the program itself. We have observed basic game-playing ability generalize from one game to another in the curriculum. We have also witnessed the transfer of rule-learning from an easier to a harder level of the same game. The ability to deal with a specific communication task (e.g., Bingo) generalizes likewise to other tasks which have elements in common with the original (e.g., Go Fish). At the level of language usage, individual children were observed to generalize some utterance forms and expressive styles across different games. These included question-asking strategies and information-getting routines, as well as instances of superordinate categorization (e.g., "is it a plant?" to replace "is it the one with green leaves, and roots and ...?") after such usage had been modeled by the teacher.

The message that these kinds of observations convey to us is that there are many different kinds of learning occurring at many different levels within this type of oral language program. Some of these learning effects are more subtle than others, but, over the long term at least, most of them can be observed in a subjective, informal way among individual children. They are not all directly measurable, but accurate description and recording of their occurrence can be a first step in the direction of more systematic analysis.

d. The general applicability of the program and the issue of classroom management

Because of its very nature, talking is noisy and cannot be carried on very meaningfully in large groups. And listening, if it is to be active and purposeful, must be sustained and concentrated, which requires that it be focused. When speaking and listening are to be stimulated and fostered in the classroom for the sake of language learning, then the special nature of these activities requires that certain changes be made in the organization and operation of the classroom. Since our training program was carried out in specially prepared, experimentally controlled environments, the question will naturally be raised as to how easily or readily our program will translate into regular, everyday classroom practice. From our own experience and the experience of other teachers and school who have tried to apply our program on an ad hoc basis, a number of suggestions can be made.

We have previously discussed how this kind of oral language program dictates new roles for child and teacher. The child engages more actively in dialogue and plays a variety of speaker-listener roles to solve communication problems with others. The teacher plays a less didactic role, becomes a more conscious language model and attempts to manage the resources of the classroom in such a way as to create situations where language learning will occur. While this requires a good deal of curriculum development on the teacher's part, it also creates serious problems of classroom management. An oral language program like ours is based on child-centered, small group, talkative activities, but how can this be introduced and accommodated to the realities of classes containing 25 or more students?

From our experience in implementing the program and in having to modify it en route, we learned that the groups have to be small to start with (perhaps only 2-3 players) in order to be manageable but, more importantly, in order to ensure that real oral practice was taking place. Given the training goals of the program, it is self-defeating to have groups that are so large that any one child gets a turn to speak or to listen actively only every 7 or 8 minutes or so. Over time one can work up to larger and more independent groups, keeping in mind that there is an optimal ratio of group size to participation time, depending on the kind of communication game being played.

The main problems of classroom management occur in the early stages of the program when the teacher is trying to introduce it into the regular curriculum. We can offer some specific suggestions for coping with this. Quite apart from the language learning goals, the management goal is to cumulatively enlarge the "game " repertoire of each and every pupil. Ideally one would like to see 100% of the students master 100% of the communication games in the curriculum. First off, the children have to be brought up to game-playing level, as it were, and to do this the teacher has to demonstrate the rules, procedures and instructions for each game. There are two strategies for doing this as efficiently and as economically as possible, and both can probably be pursued simultaneously. One strategy is to begin with simple games that can be played in dyads. If the teacher were to model the playing of a basic game with one child while the rest of the class observed, the latter could then pair off to play the game on their own. The other strategy would probably expand the class's game repertoire even more speedily. The teacher should begin by working with some of the brighter, more competent children, a few at a time, while the rest of the class continues with other assignments, like reading or writing. After learning some of the games with the teacher, this initial cadre of children can split up and begin to work individually as peer models, teaching the games to newer children. In this way the teacher multiplies her game-teaching resources and assures the progressive expansion of the oral curriculum throughout the class.[1]

Finally, there is a parallel issue that bears directly on the problem of classroom management of the games. This is the question of how to order the games in a hierarchy of difficulty so that they might be introduced to the students in a systematic sequence. The present study achieved only partial solutions to this problem and, in the games presented in Appendix II.A., we found it easier to order the levels of difficulty within a game than to arrange the different games themselves in any final sequence of complexity. More research is needed to determine which games should precede (or succeed) others over the year's curriculum and further practical experimentation with the program by teachers in their classrooms could provide valuable feedback for this issue.

---

[1] We might label this approach the "Paul Jones Strategy" after a dance common in Britain. In order to warm up a dance crowd in the slow moving early stages of an evening, the band leader would announce a "Paul Jones" which required those couples already on the dance floor to split up and bring new partners onto the floor. In no time at all the ice was broken and everyone was up and dancing!

d. Monitoring and evaluating a training program -- short- vs. long-term goals.

Evaluation of training effects can be served in different ways at different points in time. When this kind of a communication training program plays a regular and substantial role in the language arts curriculum throughout the school year, then the long-term measurement goals should probably be met through the use of a systematic, controlled testing procedure similar to the battery of tests used in the present study. Admittedly these kinds of experimental evaluation instruments are lengthy and time-consuming to administer, while the analysis of the response data can be complicated. In their present form our test procedures may not fit well with the measurement needs or testing resources of many teachers. What is needed is a boiled-down version of the competence battery that is easier to administer and score. In some research in progress, Hudson (1978) has been experimenting with the use of a mini-version of our tests in schools in Cambridge, Massachusetts. More research and development is needed to refine evaluation techniques for the measurement of speaking and listening competencies. In the meantime, given the newness of our competence battery and the lack of any normative sampling data, our measurement techniques provide the teacher with a form of criterion-referenced testing. Since the battery was developed to detect different levels of communicative ability among groups of children the teacher can use it to compare one child's current level of performance (e.g., at the end of the year) with his or her prior level of functioning (e.g., at the start of the year) to determine if any progress or development has occurred. The teacher can also use the test data to compare that child's performance with that of his classmates. The important thing to keep in mind here is the relativity of such evaluations. A child who is rated Low competent in one school system relative to his peers there might possibly receive a rating of Moderate or even High competent in another school district when compared with the levels of performance of the students in that other town. Notwithstanding this caution, within the limits and constraints of the test sample, the data produced by this kind of evaluation still have good diagnostic and prescriptive value for the teacher.

Objective measurement of communicative competence over the long term might be regarded as a form of "external" measurement of a program and its effectiveness. That is, the utility of the program's content is evaluated by means that are independent of and outside the program itself. Some warnings need to be given to the teacher to guide her use and interpretation of such evaluations. First of all, whether the teacher uses the assessment techniques developed in this study or chooses some other form of criterion-referenced testing, she must be sure that there is an appropriate correspondence between the content of the test instruments and the content of the oral curriculum. A mismatch between what is taught and what is tested could vitiate the effects of the whole evaluation program. The teacher should also be aware that there is an important distinction to be made, as we tried to point out in our discussion of generalization in (c) above, between testing that measures changes in verbal responses and testing that measures changes in underlying skills or processes. It would appear to us that the latter kind of change ought to be the more enduring and have wider implications for the child's language development. As such this type of change ought to be an appropriate object of study in a long-term evaluation program.

The second precaution, however, requires the teacher to be able to distinguish those changes in communicative competence which are directly related to the oral language program from those that are due to natural development. As will be reported in Part III of this report, even the control Ss in our study showed improvements in their test performance over time, and this could be interpreted as the effect of normal maturation and development of their communicative ability. Of course the rigors of experimental design cannot be expected in the ordinary classroom but, even without control Ss, the teacher should try to work with measurement techniques that can discriminate training from non-training effects. Thirdly, the long-term measurement goals for speaking and listening should be more integrated with the evaluation of other language arts areas like reading and writing. If, as the Bullock Report (1975) argues, the treatment of these four competencies should be more interrelated and interdependent in the realm of teaching and curriculum development, then it also makes sense to integrate evaluations in one domain (e.g., oral vocabulary development) with those in corresponding areas (e.g., reading vocabulary development). An integrated evaluation program could yield valuable intercorrelations of scores from various measures of reading, writing, speaking and listening abilities.

The kind of objective, "external" program evaluation that we have described here may not be everyone's cup of tea, for whatever reason, and many elementary teachers are put off by measurement techniques that appear to be "too psychometric". To counter this reaction we think we can offer an alternative evaluation strategy which requires a more "internal" kind of monitoring of learning effects. This approach is closely tied to the use of games and to the issues of curriculum development that were discussed earlier. Internal, within-program measurement of a child's progress is made possible when the oral language curriculum content is organized and ordered in a hierarchical fashion. This means that each game (or category of games) has a known level of difficulty so that all the games can be arranged along a continuum of complexity. Further, since each game itself can be played in easier or harder versions, a further sub-division of levels of difficulty is made possible within all of the games. An analogy can be made here with the different primer levels in elementary readers where a child's reading achievement can be measured by his progress through the different books. In a similar way, where an explicit, or even an implicit hierarchy of difficulty exists in an oral language curriculum, it becomes relatively easy for a teacher to monitor a child's progress by observing and recording his mastery of the different levels within a communication game and his graduation from one game to another. Approached in this way, program evaluation becomes a built-in part of the curriculum and year-end assessments of achievement consist of simply demonstrating the rate and extent of a child's progress. The curriculum developed for this study could be regarded as the prototype for a hierarchically organized oral language program, but a great deal of work remains to be done on developing the curriculum and refining its organization.

Fine-tuning of the program should probably also be the goal of short-term evaluation strategies. Unlike long-term evaluations, which are presumed to be more definitive and conclusive, these periodic, interim assessments can be more informal, relying, say, on weekly or monthly observations and descriptions of the child's language behavior and the quality of the match-up between a particular activity and the child's learning requirements. Tough's (1977) program could provide helpful guidelines for this sort of on-going, clinical-analytic evaluation. Short-term monitoring will be most effective when the feedback it provides at regular intervals is used to make adjustments and corrections in the day-to-day operation of the curriculum. When the more internal form of long-term evaluation of learning is preferred, then regular short-term monitoring of program operation is absolutely essential, first to establish the necessary criterion levels of difficulty within the curriculum and second, to make these levels more precise and reliable.

PART III

Assessment of Training Effects

PART III

Assessment of Training Effects

I. POST TRAINING EVALUATIONS

1. Testing Procedures

i) Administration

The same battery of primary measures used in the pretest was administered to the subjects at the end of the training period. During a three-week period towards the end of the school year, three experimenters gave the battery of perceptual, cognitive and linguistic tests to the total sample of 97 Ss. Although the core experimental grouping (EXPGROUP) contains only the 49 Low Ss, the total sample was evaluated to allow us to measure relative changes in competency status (COMPSTAT) also. The 97 Ss were also re-tested with the Stanford-Binet Vocabulary scale, and received year-end reading achievement tests too, administered by the schools as part of their annual, system-wide evaluation program.

ii) Order effects and test pairings

As before, the perceptual and cognitive tasks were administered individually to each child by the adult experimenter. However, for the verbal-communicative tasks, the subjects were again tested in pairs, in child-child dyads. This necessitated controlling and counter-balancing a number of variables affecting the sequencing of the tasks, the status of the partners, and the starting order for each member of a test pair. In order to ensure that the post-training evaluation was a real test of the generalisation of learning acquired during training, we attempted to co-vary a number of measurement variables and subject variables simultaneously. First, within the verbal domain that we have called Spatio-Temporal Language tasks (STL), we reversed the order of testing such that Ss who had gone first in the pretest now went second, and vice versa. This meant, in effect, that Ss who had received the 'A' version of the test items in the pretest now received the 'B' version, and vice versa. Similarly, in the verbal domain that we have called Logical Language tasks (LL), the sequence in which the tasks were attempted was also changed. A subject who had had Encoderposition A in the pretest now took position B, and vice versa. This too meant, in effect, that Ss who had started first and encoded the 'A' targets in the pretest, now started second and encoded the 'B' targets, and vice versa. Next, we attempted to pair Low Ss with partners who came from a classroom other than their own. We further tried to reduce familiarity effects by requiring that the Low Ss be tested with children who were not their training group partners.

On top of all this the subjects were also required to change partners when they moved from the STL to the LL group of tasks. Within the limits imposed by all of the preceding constraints, we attempted to control one final variable, the status match-up of the test partners. It will be remembered from Part I of this report that, since the competency status of each S was unknown prior to the pre-test evaluations, the nature of the pairings in the testing dyads could not be controlled at that time. As a consequence, the COMPSTAT of the dyads was not randomly distributed and instead, as Table 1-a indicated, the distribution of COMPSTAT match-ups was somewhat lopsided. In the post-test, the constraints of order and position likewise prevented an even distribution of COMPSTAT combinations across all the dyads tested, but we achieved some limited success in making the distribution less skewed than before (Table 1-b). As Table 1-b indicates, we were more successful in some schools than in others in pairing Low Ss with an equal number of Low, Moderate or High competent partners. The continued preponderance of L/M pairings here (i.e., a Low competent child paired with a Moderate competent child) can probably be attributed to the greater number (and hence availability) of Moderate Ss in the sample.

2. Data Management

i) Transcription and data reduction

After the testing was completed, the next two months were spent transcribing, scoring and coding the data. The perceptual and cognitive data were scored right off the test protocols, but the verbal-communicative data had to be transcribed from the cassette tapes on which they had been recorded during testing. It took a team of transcribers a full month to transcribe each of the verbal communicative test sessions for all 97 Ss. This produced hundreds of pages of transcript, which were thereafter re-checked and verified by independent raters in the succeeding weeks.

ii) Scoring and data processing

The rest of the summer was spent scoring and preparing the data for computer processing. (The scoring scheme is as described in Part I.) Since some of the perceptual, STL and LL measures had undergone redefinition or reformulation since the pre-training evaluations, this meant that some of these pre-test data had to be re-scored also along with the post-test measures. The very lengthy process of scoring, checking, coding, and keypunching of data eventually produced 6 cards' worth of data for each of the 97 Ss. Each test produced Item Scores (within tasks), Task Scores (within domains) and scores by Domain for the perceptual, cognitive and two linguistic areas (STL and LL). Computer analyses of these scores and their interrelationships generated over 400 variables per child.

The next task was to reduce this mammoth pile of data to manageable proportions and to identify appropriate descriptor and analytic variables for measuring the study's effects. The main computer programs used for this purpose were SPSS and a MANOVA program called NYBMUL, based on the work of Bock and Haggard (1968) and Finn (1974).

iii) Gain scores, covariates or unadjusted post-test scores?

With the traditional kind of pre-test/post-test, experimental vs. control group design used in this study, decisions had to be made about the analysis strategies that should be employed to measure treatment effects. Recent discussions of this issue in the research literature (Bock, 1975; Huck & McLean, 1975) have compared the relative merits of gain scores, covariate scores, and unadjusted scores. Each of these approaches were considered for measuring post-training changes within the context of our particular experimental contrasts. At one time or another during the post-test evaluations, a particular analysis would dictate that one or another of the 3 analysis strategies should be used. For the most part, however, unadjusted post-test scores or gain scores were the two most frequently used forms, with covariance analysis used less frequently. The reliance on unadjusted scores needs to be explained, but in order to understand this decision the relative merits of the three strategies need to be considered first.

Each of these approaches is concerned with measuring change in some absolute or relative way. Each leads to the formulation of some kind of comparative statement about treatment effects. Yet, subtle though the distinctions might seem, each approach poses its question in a slightly different way, and is based on different assumptions. We might say, for instance, that the gain score approach is concerned with the amount of change: that is, when pre-post gains are calculated as "net difference" scores they are less concerned with differences in starting levels and focus more on the average gains within particular groups. Covariance analysis (where the covariates are the pre-test scores), on the other hand, might be viewed as addressing itself to the extent of change, where the interpretation of treatment effects is couched in terms of "more or less" rather than "how much." Covariance thus provides a more relative measure of change which attempts to assess the comparative gain between two groups when differences in starting levels are taken into account.

The other way to take account of differences in starting levels is to demonstrate at the outset that there is initial equivalence between the groups, as a precursor to using unadjusted post-test scores to measure differences in absolute levels of post-test performance. This latter strategy was the primary one used in the present analyses and its use can be justified on a number of grounds.

For one thing, it appeared to be the simplest and most parsimonious solution, given the kinds of controls and precautions that had been taken to equate the pre-test scores of the treatment groups. In advocating parsimonious research analyses, Huch and McLean state that "in terms of data analysis, it appears as though many researchers have a tendency to do things in a complicated manner even when a simple procedure can accomplish the same goal" (1975,p. 516).

A second condition that contributed to our use of raw post-test scores was the manner in which the experimental groups were formed and their initial equivalence established prior to training. It will be remembered that all Ss had gone through an evaluation and selection procedure which had differentiated and grouped them according to criteria of competence before any treatment was applied. However, it was not enough to be satisfied with the accuracy of the categorisation of the Low Ss even though the trichotomising of competency status (COMPSTAT) appears to have been reliably demonstrated across groups by the pre-test data (Part I). We also had to deal with variability of scores within-groups (EXPGROUPS) in order to ensure that what made the difference in EXPGROUPS I or II or III was the treatment, and not the fact that all the Low Ss in one group had higher or lower pre-test scores than the Low Ss in one of the other groups. We addressed ourselves to this issue of pre-test variability within the four Low EXPGROUPS in two ways. First, by assigning Low Ss randomly to each of the four experimental groupings and second, by double-checking that these groups were indeed relatively equivalent by means of univariate analyses of their levels of performance on the pre-test measures.

While we naturally encountered individual variability, our tests of initial equivalence (Part I) satisfactorily demonstrated that, on the average, and at the level of the groups, there were no significant differences in virtually all of the pre-test measures between the four Low groups. Given this prior condition, therefore, we chose to work with the unadjusted post-test scores to measure if, on the average, children who received our communication training improved their competency relative to children without such treatment.

Neither the gain score nor the covariate approach necessarily have to prove or disprove initial equivalence between groups. Another difference between these approaches and the use of unadjusted scores is that both the gain score and covariance approaches are often used in open-ended studies which contain no prediction about the occurrence or direction of change (if any) between grossly differentiated groups. Our use of unadjusted post-test scores as a primary analysis strategy, however, had more specific purposes than this. Premised on initial equivalence, *a priori* hypotheses and experimental contrasts were constructed concerning the performance of the three treatment groups together and separately. As expressed in the original proposal, the first hypothesis was couched in a conservative way, *viz.*, that there *would* be a positive effect of being in a training group. No additional conjecture was made about the anticipated amount or relative extent of that effect and the choice of analyses via unadjusted post-test scores reflects that simpler, conservative position. The remaining experimental hypotheses do deal with more relative differences in performance within the three training groups I, II, and III. Here, to be consistent, raw post-test scores were also the basis of many analyses, but other exploratory analyses using gain scores were also employed to assess these within-group Pairing effects.

Finally, one other condition gave us confidence that unadjusted post-test scores could be reliably sensitive to treatment effects. This constituted establishing rigorous controls over test-instrument effects. Despite establishing initial equivalence, there was still a small amount of between-group variability found at the pre-test level which seemed to be attributable to instrumentation effects. Such effects (often confounded with test-item learning and familiarity effects) might interfere with the interpretation of treatment effects. Our awareness of these potential instrumentation effects enabled us to deal with them during the post-training testing sessions by counterbalancing and controlling such factors as test-position, test-partner, test-order and test-form. We feel that the latter controls further contributed to increase the reliability of the unadjusted post-test scores as suitable indices of communication training effects.

II. RESULTS FOR PRIMARY MEASURES: EXPLORATORY ANALYSES

As stated in the original proposal, these results will deal only with the data for the target sample of 49 Ss who made up the four Experimental and Control groups. Descriptive analyses will be presented first for each of the Perceptual, Cognitive, and Linguistic domains, and these will attempt to pinpoint any trends in the data from pre-test to post-test. This will be followed by a confirmatory analyses where a fuller statistical evaluation will be made of the significance of any changes in the major analytic measures. Finally, an analysis will be made of the remaining secondary measures of reading and vocabulary.

1. Perceptual Domain

As before, the MFF test results were analysed with 9 measures of speed, accuracy/inaccuracy, and speed x accuracy relations. The bulk of the analyses concentrated on 5 of these measures, as presented below. A brief comment can be made about the order of difficulty of the perceptual test items. The same trend as reported for the pre-test data was also observed in the post-test results, namely, that the order of difficulty very closely approximates the order of presentation of the items (1 through 5). As in the pre-test, the 49 Ss taken as a whole still spent more time on the average with the later (and harder) items, and made more errors on the harder items. The noticeable change from pre- to post-test is that the Ss take a little more time on each item and make relatively fewer errors.

The significance of these changes can best be examined at the level of the within-groups and across-task data, and this is presented next.

i) Accuracy

The accuracy of the first choices improves for all groups after training (Table 2). Since these data represent the unadjusted post-test scores, it is necessary to compare the absolute levels of post-test performance across the four groups. This indicates that the Experimental Ss as a whole do slightly better than the Control group, but the differences between-groups is not significant (F = 0.511, p = 0.677).

Within the Experimental groups, group II (Lo/Mod) gets the most correct, and it is interesting, in this respect, to note that they also slow down more when making their choices (Table 3). The Lo/Lo Ss (group I) make the least number of correct first choices of all groups.

In relative terms, all groups improve from pre to post, with the Control group having a slight edge over the others.

ii) Speed

There are no clear-cut trends or significant results in this measure of the average time taken on each item to make a choice (Table 3). In absolute terms, the Control group is noticeably slower than the Experimental groups, whether taken together or separately, while two of the Experimental groups (I and III) are noticeably faster than the others. There are no significant between-groups differences (F = 0.242, p = 0.867).

In terms of pre-post change, this is negligible except for the Lo/Lo Ss (I) who became faster, and the Lo/Mod Ss (II) who became slower in making their first discrimination, on the average. (It should perhaps also be taken into consideration that this speed measure is averaged across the 5 items of the task, so that it does not necessarily take account of the increasing order of difficulty of the items.)

iii) Accuracy x speed

Tables 4 and 5 show the relations between time-to-choice and the accuracy or inaccuracy of the choice.

There are no clear-cut trends in the post-training accuracy scores, except that all the groups slow down, i.e., they take more time to make a correct first choice (Table 4). Although some groups (e.g., II and IV)are slower than others, none of these between-group differences are significant (F = 0.378, p = 0.769). In absolute terms the Control Ss are noticeably slower then the Experimental Ss in making their correct choices, while within the Experimental groups it is again the Lo/Mod Ss (II) who are the slowest. In relative terms, however, the most dramatic change from pre to post occurs in the Lo/Hi group (III) who slow down considerably more than anyone else.

In general, the results for Inaccurate Choices (Table 5) is the reverse of the above. That is, on the average, incorrect choices are made faster after training, and they are made more quickly, on the average, than correct choices. The Control group is faster than the Experimental groups, but these group differences are not significant (F = 0.639, p = 0.594). The Lo/Mod experimental group (II) produced some results that ran counter to the general trend. They are slower than any of the other groups in the post-test in making their incorrect choices and they are the only group who become slower for such choices after training than before.

iv) Accuracy ratio

This composite measure of overall accuracy of perceptual discrimination shows all groups improving their general accuracy over time (Table 6). After training, the scores of the Experimental groups are higher than the Control group, but none of these differences is significant (F = 0.516, p = o.674). Within the Experimental groups the Lo/Mod Ss (II) once again have the highest accuracy scores. Regarding relative change from pre to post, the largest amount occurs among the Control Ss.

v) Correlations among measures

Pearson correlations for the 5 measures were run on all 49 Ss as one group in the post-test analyses. Although it could be argued that this might confound the results with the training effects, it was preferable to calculate the correlations for the whole group in this way in order to be consistent and comparable with the pre-test correlations for the 49 Low Ss. Moreover, since the other results presented above do not indicate any clear-cut effects of training on perceptual performance, the disadvantages of lumping experimental and control Ss together in the same correlational analysis are minimised.

As in the pre-test, the r's are still significant, but the strength of the relations is only low to moderate (Table 7). After training, Speed and Accuracy now seem to have a close relationship which seems largely due to Speed being more controlled in the interest of Accuracy. That is, on the average, a longer choice-time tends to be more accurate. The increasing importance of speed is also seen in the near perfect correlation of this measure with the composite measure of time-to-correct-choice. With the exception of this correlation, this latter measure weakens its relations all across the board. Speed also increases its relation with the composite Accuracy Ratio measure -- a measure that has virtually a synonymous relation with straight Accuracy.

Comparison of the pre-test with post-test correlations further strengthens the impression that Accuracy of choice is affected by changes in Speed of choice. The pre-test measure of Accuracy best predicts itself in the post-test analyses, as well as the post-test Accuracy Ratio measure. As such, it could easily be substituted for this Ratio measure which turns out to be a poor predictor of itself (r = .25) from pre to post.

The pre-test measure of Speed, rather than Accuracy, has some of the better predictor relations with the post-test measures. Its relatively stronger relation with the post-test composite measure of time-to-incorrect choice suggests that if a subject was fast in the pre-test he might have a lower mean time (= faster) when making an incorrect choice in the post-test. This kind of "more-haste-less-correct" result can also be found in the relation between Speed x Inaccuracy in the pre-test with Accuracy in the post-test (r = .3179, which is this pre-test measure's best predictive relationship).

This suggests that when a child makes his incorrect choices more quickly (= a lower score) in the pre-test there is some tendency for him to have a lower number of correct choices in the post-training evaluations. Conversely if he is slower (= a higher score) in making erroneous choices before training there is a likelihood that he will make a higher number of correct identifications after the training period is over.

In general then, and without consideration of experimental status, slowing down appears to be a positive condition which contributes to a child's accuracy of discrimination. The extent to which this speed and accuracy relation can be influenced by exposure to a communication training program remains to be demonstrated more clearly by further research.

vi) Summary and discussion of the perceptual results

There appear to be some very general trends in the perceptual discrimination data, but there do not appear to be any systematic or significant effects of communication training on this domain. On the whole, all the Ss tend to be a little more accurate in making their first choice of target picture and to go about this with slightly less haste in the post-training evaluations. There do not, however, appear to be many experimental effects that would warrant further analysis. There is a tendency for the Experimental groups to be a little more accurate after training than the Control group, but not necessarily slower than them.

A more interesting finding comes from studying the effects of pairing on perceptual performance. The most noticeable and consistent results occur among the Low Ss who are paired with partners one step above them on the competency scale. These Lo/Mod Ss (group II) appear to have adopted a general strategy of "less haste" after training, and associated with this is a higher degree of accuracy than that found in the other Experimental groups. In looking for an explanation of this result, an examiniation of the data for the High competent Ss (N = 24) proved to be instructive. These High Ss, both at pre- and post-test, are markedly slower and more accurate overall in making their discriminations than any other group. In this regard, therefore, the one group among the Experimental Ss who come to resemble the High competent Ss is the Lo/Mod group after training. The data from this part of the study, therefore, suggest that for the perceptual component of communicative competence at least, it might be advantageous to have less able children work with peer models who are more advanced (but not too much more advanced) than they are.

As far as the utility and value of the analytic measures is concerned, the performance of the composite Accuracy Ratio measure is disappointing, for it adds little beyond what the simpler Accuracy measure can tell us. More interesting and discriminating results are provided by the measures of Speed and Speed-with-Accuracy which indicate that time is an important factor in a perceptual discrimination task like this. It is a variable whose relations with Accuracy should be explored further. It would appear, in retrospect, that the composite Accuracy Ratio score should have been revised to include some expression of Speed in its formula. In such a revised form it may have been more sensitive to the effects of training.

2. Cognitive Domain

The influence of the verbal training program on the Low Ss's egocentricity and perspective-taking ability was examined in a variety of complementary ways. The results will be discussed separately for the Tasks, the Measures, the Summary Measure (which combines all the scores from all the tasks), and the interrelationships between these three categories.

1) Tasks

The post-test results for the three cognitive tasks are shown in Tables 8, 9, and 10.

The one-item task, House Construction (with a maximum possible score of 3) shows some clear trends for the positive effects of training, with the Experimental groups scoring higher than the Control group in the post-test (Table 8). However, these apparent group differences do not reach statistical significance (F = 2.16, p <.10). There also appears to be an ordered Pairing effect within the three Training groups (III > II > I) with the Lo/Hi and Lo/Mod Ss scoring higher than the Lo/Lo group. A test of this contrast (groups II + III vs. I) proved to be significant (t[45] = -2.22, p <.03).

In relative terms, the Lo/Lo group I show the same, small amount of pre to post change as the Control group, while the greatest gains are made by groups III and II respectively. In t-test comparisons of pre vs. post scores for each of the four groups, only the changes for the Lo/Mods (t = 3.0, p <.01, df 11) and the Lo/Hi's (t = -1.61, p <.069, df 10, one-tailed) reach statistical significance.

In both absolute and relative terms, therefore, Expgroups II and III are the strongest performers overall.

The results for the House-Pictures task (maximum possible score = 12) are similar to those for House Construction (Table 9). Once again the post-test scores are greater for the Experimental Ss than for the Controls, and there is an ordered increase in scores across the three Training groups attesting to the differential effects of Pairing. Despite these indications, the group differences in Table 9 are not statistically significant (F = 0.547, p = .653), nor do any of the within-group contrasts reach significance. In terms of relative change from pre to post, the Experimental groups together gain more (+0.72) than the Controls (+0.31). Individually, the Lo/Lo Ss again resemble the Controls, but the increase in score for groups II and III is both noticeable and significant (respectively, t = 2.11, p <.029, and t = 2.17, p <.027, one-tailed). Once again the Lo/Mod and Lo/Hi Ss show the strongest effects of training.

The Piagetian Mountain Perspective Task produced results to attest to the significant effect on cognitive functioning of participating in the training program (Table 10). Once again there are clear-cut trends for Training and Pairing effects. All three of the Experimental groups score significantly higher than the Control group (F = 3.20, p <.03, df 3/45). There is an ordered pattern of performance (III > II > I) showing the beneficial role of modelling within the training groups. Contrasts (using pooled variance estimates) gave significant results for comparisons of groups II + III vs. I (t[45] = -2.67, p <.01). An additional comparison of the largest performance difference — that between the Lo/Hi Ss and the Controls — produced a further significant result (Dunnett's t[45] = 2.37, p <.04).

In relative terms, the performance of the Control Ss actually deteriorated over time (-0.15) even though they maintained the same level of variability from pre- to post-test. The gain for the Experimental Ss, by comparison, was strongly positive (+0.89). Individually, the largest and most significant score gains from pre- to post-training were made by the Lo/Mod group II ($t = 2.77$, $p < .009$, df 11, one-tailed) and the Lo/Hi group III ($t = -1.94$, $p < .04$, df 10, one-tailed). These results are consistent with those from the other two tasks.

As with the pre-training data, it is possible to use the TASK data to see if there is still a hierarchy of difficulty across the three tasks. As measured by percent of maximum possible score achieved in each task, on the average (for N = 49), the same order of difficulty is observed in the post-test data. The House Pictures task is the easiest (with 75.5% of maximum) followed by Mountain Perspectives (61% of maximum) and House Construction (the hardest task, with 57% of maximum achieved). The level of achievement goes up for all tasks from pre to post, with the largest gain (+12%) being made on the hardest task, House Construction. Within the Experimental groups an equal amount of gain is made on House Construction and Mountain Perspectives.

ii) Scores

The three-point scoring system (1 = Egocentric, 2 = Intermediate, 3 = Non-Egocentric) provides a measure of any changes in egocentric perspective, across tasks, after participation in the training program. Tables 11, 12, and 13 indicate that there are clear-cut trends and some significant results in the data provided by these measures of the eight item scores derived from the tasks.

At the egocentric end of the scale (the "1" scores) the Experimental Ss and Controls have moved in different directions in the post-test period. The Experimental Ss now produce fewer "1" scores while the Control Ss produce more (Table 11). In absolute terms, the Experimental groups together have lower post-test scores (= less egocentricity) than the Control group, but a univariate analysis of the between-group differences in Table 11 did not prove significant. Within the training groups, II and III have fewer "1" scores than group I, but the Pairing effects are less well-defined here, and none of the statistical comparisons proved to be significant.

Looking at the change from pre- to post-training, all of the treatment groups reduce their egocentric choices after training ($\bar{x}$ gain = -0.72) while the Control group moves in the opposite direction ($\bar{x}$ gain = +0.77). All three of the Training groups reduce their egocentric choices from pre- to post-test, with group II (Lo/Mod) showing the greatest reduction (-1.33).

The intermediate "2" scores also decrease at the post-test for all groups except the Lo/Lo Ss (Table 12), who increase theirs by a moderate amount. The Training groups together have higher post-test scores here than the Control group, but none of these differences is significant ($F = 1.62$, $p < .197$, df 3/45). Within the Training groups, the difference between group I and groups II + III is significant ($t = 2.02$, $p < .056$, df 45), showing that the Lo/Mods and Lo/Hi's are effectively reducing their "2" scores after training.

Over time the largest change in "2" scores is by the Control group (-1.23), followed by the Experimental Ss, and groups II and III.

The most dramatic evidence for the positive effects of communication training on cognitive perspective-taking comes from the data for the "3" scores, the measure of non-egocentricity (Table 13). These show clear and predicted trends for both Training effects and Pairing effects. The three Training groups together have a higher number of "3" scores after training than the Control group, and the between-group differences in Table 13 are significant ($F = 2.72$, $p < .055$). There is also an ordered effect of Pairing in the post-test scores of the three treatment groups in Table 13, with III > II > II, as predicted. The Lo/Hi's and Lo/Mods perform here at a level twice as high as that of the Lo/Los and a test of the contrast groups II + III vs. I proved to be highly significant ($t[45] = -3.03$, $p < .006$).

In terms of relative changes, the Lo/Mod and Lo/Hi groups also show the largest shift towards non-egocentric perspective after training. Both the Control *Ss* (who were among the highest scorers in the pre-test) and the Lo/Lo *Ss* show comparatively little change. The gain score analysis for the training groups together shows a pattern similar to that provided by the post-test scores above. The Experimental groups show an average gain after training (+1.19) that is 2 1/2 times that made by the Control group (+0.46).

An overview of the results for this 3-point scale of egocentricity indicates definite Training and Pairing effects benefiting the Experimental groups. These *Ss* become less egocentric than the Control *Ss*, while within the Training groups there is an ordered pattern of performance (*viz.*, III > II > I) indicating that the amount of change for each group of Low *Ss* is related to the level of competence of their peer models. The nature and extent of the changes differs for each of the four groups. The two most consistent performers, both absolutely and relatively, are groups II and III, who dramatically increase their "3" scores (non-egocentric) while decreasing their "1" and "2" scores. The effect of Pairing is less extensive for the Lo/Lo *Ss*, who move towards a greater preponderance of the intermediate "2" scores, suggesting, perhaps, that egocentricity is not a dichotomous, either-or dimension, but that there are transitional stages in the decentering of perspective.

Finally, there is no consistent pattern for the Control *Ss* over time. They decrease their high beginning-level of "2" scores, but fully 63% of this score change goes to contribute regressively to an increase in the egocentric "1" scores.

iii) Summary measure, COGSUM

The correlationships (see below) between the tasks indicate a substantial amount of independence between the sub-tasks, while the 3-point scoring system appears to be a solidly discrete scale with good discrimination power. In view of these factors, the summary measure, COGSUM (maximum possible score = 24), which sums the scores (1 - 3) across all items (8), proved to be an economical, representative, and valuable single measure of ability in this dognitive domain (Table 14). COGSUM demonstrated significant results in all the major categories of analysis. The post-training scores (Table 14) show clear-cut, positive trends for both the Treatment and Pairing conditions. The three training groups together make more non-egocentric choices than the Control group, and the between-group differences in Table 13 are significant ($F = 3.13$, $p < .04$, df 3/45). There is an ordered Pairing effect, as predicted, within the Experimental groups, with III > II > I. A test of the contrast, groups II + III vs. I, is significant ($t[45] = -2.66$, $p < .01$). Comparisons of groups II vs. III and II vs. I were not significant. Additional post hoc comparisons via Dunnett's t-test showed that the highest performers, the Lo/Hi *Ss*, were significantly different from the lowest performers, the Lo/Lo *Ss* and the Control *Ss* (III vs. I, $t[45] = 2.75$, $p < .02$; III vs. IV, $t[45] = 2.32$, $p < .05$).

In relative terms, the average gain after training (+2.11) for all the Experimental Ss together is almost seven times greater than the change from pre to post shown by the Control Ss (+0.31). Within the Training groups, the greatest gains are made by the Lo/Mod and Lo/Hi Ss, repeating the pattern reported above.

iv) Inter-correlations of Tasks and Measures

Table 15 shows the pooled, within-group correlations for all the Cognitive Tasks, Measures, and Summary Measure, in the post-test period, for N = 49.

For the Tasks, all correlations, as in the pre-test, are still only low to moderate. Relations between Tasks increased a little, especially between House-Pictures and House-Construction. In so far as each task predicting itself is concerned, the two easier tasks (House-Pictures and Mountain-Perspective) do this best.

For the Measures, the post-test correlations data show that the '3-score' continues to be the most discriminating measure of cognitive achievement. The 3-score (non-egocentric) maintains its inverse relation with the 1-score and 2-score, markedly doubling its correlation with the latter (r = -0.629). From pre-test to post-test, each score predicts itself with fair (2-score) to moderate (3- and 1-score) success, confirming the observation from other results that the major effect of training is the shift from egocentric (1-score) to non-egocentric (3-score) perspective. The 3-score especially is a fairly good predictor of success from pre to post (r = 0.51).

The same degree and shape of Task x Score relations is maintained in the post-test results, i.e., the relation follows the hierarchy of Task difficulty and the scale of egocentricity. The easiest task (House-Pictures) strengthens its negative relations with the 1-score (r = -0.612), while the other two tasks increase their positive relations with the 3-score.

Once again, in view of the continued, relative independence of Tasks and Measures, the best overall measure of post-test achievement in this perspective-taking domain is the SUMMARY Measure, COGSUM. The solid COGSUM x Task relation increases in the post-test evaluations, with the degree of this relationship conforming, once again, to the hierarchy of task difficulty. Similarly, the positive and negative COGSUM x Scores relations increase from pre to post, with its relation with the '3-score' now virtually one-to-one (COGSUM x '1-score', r = -.76, COGSUM x '3-score', r = . 93).

Finally, a Reliability analysis was also made of the three tasks in the Cognitive scale, for N = 49, and the results can be reported briefly here. While the usual problems exist when post-test responses from Experimental and Control groups are lumped together, the item-analyses still display general trends similar to those reported in the correlational data above. Each task retains its strong independence in the scale (the average inter-item correlation is still only .295), but at the same time increases its relations with the other tasks and with the scale as a whole. The strongest relation is between House-Construction and House-Pictures (r = .395), and both of these tasks have the strongest relations with the scale as a whole (r = .374 and r = .399 respectively). There are still significant differences between the measures, although this may be partly due to the unequal lengths of the tasks.

However, the reliability of the scale as a whole improves from pre- to post-test, with coefficient alpha rising to 0.511 (standardised item alpha = .556). The hierarchy of difficulty across the three tasks is again also confirmed by the item-total statistics. The scale mean is highest when the hardest task (House-Construction) is deleted and lowest when the easiest task (House-Pictures) is deleted.

v) Summary and Discussion of Cognitive Results

The post-training results for the cognitive component of communicative competence provide a variety of significant findings which have direct implications for the effectiveness of the language training program and the role of peer models in such a program. The results also raise fascinating issues concerning the relations between pragmatic, functional language use and cognition.

Overall, it would appear that the three perspective-taking tasks are justified as relatively independent but complementary measures of cognitive ability. They provide a representative variety of perspective-taking problems requiring an assortment of active and passive responses. The post-test data confirm that the three tasks are ordered in a hierarchy of difficulty which is sensitive to the effects of training. It is noteworthy that the hardest task is the one-item, House Construction problem which requires the most active kind of response of all. This suggests that passive comprehension of another's perspective may actually precede the active skill of concrete, overt re-construction of that perspective in ontogenesis.

The post-training data also indicate that the 3-point scoring system provides a solid scale of egocentricity, with discrete measures of ability at each level. Such a scale helps illuminate the special effects of Training and Pairing on this cognitive variable. For example, when the data of Tables 11 - 13 are taken together, they show the differential cognitive effects of the communication training program for each of the four Experimental groups. We have already pointed out how the Lo/Mod and Lo/Hi Ss significantly increase their non-egocentric '3-scores' after training, while the Lo/Lo Ss do not move as far along the scale, showing primarily an increase, after training, in transitional '2-scores'. Meanwhile, the Control group, without the benefits of training, show no systematic decrease in egocentricity.

The overall pattern of results for these measures suggest some fascinating possibilities. First, that oral language training (as conceived in this study) can directly affect cognitive functioning by influencing the direction of change, in a positive manner, along a scale of egocentric perspective-taking. Secondly, when one combines training with Peer-Modeling, one can influence not just the direction of change, but also the distance moved along the scale of egocentricity. The results for the Lo/Mod and Lo/Hi Ss directly imply that the extent of decentering by a child in the cognitive tasks is a function of the level of competence of his training partners. Thirdly, the transitional status of the Lo/Lo Ss after training further suggests that the cognitive scale is a continuous rather than dichotomous dimension. As such it would appear to be differentially sensitive to the influences of intervention and evaluation. Future developmental studies of the relation between cognitive perspective-taking and functional language training will have to take account of such variables as: the age of the child, the nature of the training program and the role of peer-models, and the use of active vs. passive measurement instruments.[1]

---

[1] There is a less direct implication in the cognitive data that rate of change in perspective-taking might also be an issue that is sensitive to language training, but the suggestion from the EXPGROUP results is that the significance of this would be better tested in a study using a larger sample, a longer training period, and more interim assessments of cognitive functioning.

In this context it is also worthwhile to observe that the Task x Score data also contribute to the picture of post-training changes. The easiest task (House-Pictures) contributes fewer '1-scores' in the post-test, while, the other two harder tasks contribute more '3-scores', suggesting that a major effect of the communication training is to move children in the direction of more positive success on the harder tasks.

The relations between the tasks and scores and the SUMMARY MEASURE confirm that COGSUM is the one best measure of ability in this domain, and this makes it a good candidate as an analytic variable for the later multivariate analyses. An interesting pattern of results is revealed in the relations between COGSUM and the most important sub-task and sub-score items, viz., Mountain-Perspective and the '3-scores'. All three of these variables produced statistically significant results for the effects of Training and Pairing. Tables 10, 13, and 15, for instance, provide the basis for an argument that a major effect of training is to increase the number of 3-scores significantly among the Experimental group, and that it is the Piagetian Mountain Perspective task in which this result has one of its stronger manifestations. The Piagetian task is an independent task that was added to our battery of Cognitive measures at the outset. Its positive relations with important measurement outcomes in the battery as a whole provide some degree of external validation of the cognitive measures and the scale of egocentricity as well as of the general cognitive value of this kind of training in oral communication.

Finally, the statistical significance of the Pairing effects in the cognitive data (with the recurring III > II > I pattern) provides a finding which appears to be original and far-reaching in its implications. These results indicate that the degree of cognitive benefit from this kind of language training for Low competent Ss is in direct proportion to the level of perspective-taking ability possessed by their peer models. Most importantly, this result and the results for treatment effects, could be taken as valid instances of generalisation of learning beyond the curriculum. That is, there are effects of participation in a language education program which are transferred to an essentially non-language, cognitive domain. We could speculate at length about the role of linguistic processes in transmitting or mediating such cognitive effects, but we will have to postpone that temptation until after an evaluation of the results from the Language domain, to which we turn next.

3. Language Domain

A. Spatio-Temporal Language (STL)

1) Summary Overview

As was done with the pre-test data, we can begin these analyses of the post-training language data by taking an overview of the basic results for all 49 Ss.

Leaving aside considerations of treatment differences for the moment, we can see from Table 16 that all of the Low Ss taken together increase their fluency and increase the quality of their message content by improving their ratio of positive to ambiguous propositions. There is a 13% increase in the number of propositions (n = 1127) generated for the instructions and directions used in the Toy Village tasks. This means that, on the average, each child now uses about 23 propositions for his three test items, and this represents a shift from approximately 7 to 8 propositions per item.

For N = 49, the proportion of positive propositions in all messages increases to 74% while there is a corresponding decrease in ambiguous propositions to 23%. The number of negative propositions is now an even tinier percentage (3%) of total message output and, as before, they will be ignored in the analyses. There are obviously interesting differences in the levels of performance and the amounts of pre/post change represented in the data of Table 16 for the four Experimental groups, and these relationships will be explored further in the analyses below of the measures of prepositional usage, and of message fluency, quality, ambiguity, efficiency, and organization.

N.B.: Position Effects

It should be noted, too, that in all the statistical analyses presented in Tables 16 - 30, there are no significant effects for any of the variables for the Position factor, Test Form A vs. Test Form B. Accordingly, equivalence of Test Forms will be assumed (unless otherwise noted) in the discussion below which will concentrate on the Training and Modeling effects.

2) Prepositional Usage

The range of post-test scores for correct usage of the six test prepositions is still from 2 to 6 (the maximum), but the group scores have all moved up towards the ceiling, so that there is little room for differentiation between the four Experimental groups (Table 17). The groups started out essentially equivalent in the pre-test with barely a half-point separating them, and an identical narrow range of score differences is evident in the post-training data in Table 17. It is not surprising, therefore, that the statistical analyses show no significant effects for Expgroup or Test Position. As far as pre-post changes are concerned, the largest gains are made by the Control Ss (+1.08) and the Lo/Lo Ss in group I (+0.77). So, in general, it would appear that there are no special benefits of training for increasing correct usage of these prepositions.

3) Quantity

The results for this measure of communicative fluency, or message length, are presented in Table 18. As Table 16 indicated, the number of prepositions per message increased after training and this increase is particularly noticeable among the three Training Groups whose post-test scores are all higher than the

Controls'. However, the analyses of variance do not confirm the statistical significance of this trend (Table 18). There is also an interesting trend for Pairing effects in the post-test scores for the three Expgroups, with Lo/Los scoring higher than Lo/Mods, who score higher than the Lo/Hi's (i.e., I > II > III).

There do not appear to be any clear-cut patterns or trends concerning the shift in scores from pre to post. The Experimental group as a whole and the Control group increase their fluency by a virtually equal amount, while within the three Expgroups, group II (Lo/Mod) increase their output by about one proposition longer than I and III.

Looking at all Ss together, it is also interesting to note that the range of scores for message length (from a minimum of 6 propositions to a maximum of 43) remains stable over time and is virtually identical to the range of pretest scores.

4) Quality

As before, the measure of message quality, in a pragmatic-linguistic sense, is simply the total number of positive propositions generated in the First Messages of all three items of this communication task.

It was noted in Part I that the amount of quality that can be expressed in a message must be constrained, in a relative sense at least, by the length of that message. However, it was also demonstrated in the pre-test data that the correlation between quantity and quality was quite strong and positive. The strength of that relationship is maintained at the post-test level (r = .70, n = 49). In view of this, and because it was shown (above) that there were no significant differences between the groups on the measure of quantity, it seems legitimate to look at propositional quality by itself, considering it as an independent measure of message content and ignoring, for the time being, variations in message length.

The results for message quality are in Table 19, and they are all in the predicted direction. The Experimental Ss, as a whole, do better than the Control group, who are the lowest of all four groups, scoring almost three points below the highest group, the Lo/Lo Ss in group I. Despite these obvious trends, however, there are no significant effects for Expgroup differences in the analyses of variance in Table 19 (F (3,41) = .985, p = .409).

There is an equally obvious trend for Pairing effects in the post-test scores of the three Lo groups in Table 19, but its direction is somewhat unexpected. There is an ordered hierarchy of performance across the groups of the form I>II>III, i.e., the Lows paired with Lows do better than Lows paired with Moderates, who do better than Lows paired with Highs. This same ordering of quality of message is also apparent in the comparison of changes in performance from pre- to post-test. Once again the Experimental Ss as a group gain more (+4.58) than the Control group (+3.77), while within the three Training groups the Lo/Los gain more (+5.5) than the Lo/Mods (+5.08), who gain more than the Lo/Hi's, who have the lowest gain of all (+2.91). Unfortunately, this systematic pattern of results did not achieve statistical significance in a oneway analysis of variance of gain scores for this measure. For the between-groups effect, F(3,45) = .807, p = .499. Further orthogonal contrasts (e.g., Experimental vs. Controls, II vs. III, I vs. II + III) using F tests and pooled variance estimates, as well as other contrasts with Dunnett's t-test, failed to produce any further significant results.

5) Efficiency

The results for the basic measure of propositional quality showed positive trends for the effects of communication training. What happens to these results when we correct the quality measure for variations in message length in order to produce a more economical measure of Message Efficiency? We will consider the three methods of calculating such an Efficiency Ratio.

Table 20 shows the results for the straightforward, untransformed Positive/Total ratio scores, expressed as a percentage. Since the pre-test analyses indicated that there was a significant lack of initial equivalence among the four groups at the outset, the post-test scores in Table 20 have largely illustrative value, since their utility for interpretation purposes is limited. These data do show that all groups increase their percentage efficiency to more than 70%. There is roughly a six-point spread of scores separating the four groups, but the group differences are not significant (F (3,41) = .375, p = .772).

Because of the lack of pre-training equivalence among the four groups, a more appropriate test of training effects is a gain-score analysis of the relative amounts of change in efficiency ratios for each group after training.

The results for this second method of calculating efficiency are illustrated by the Ratio Gain scores in Table 21. This provides a much clearer picture of what is happening to the Training groups in a manner that is consistent with the results for the Quality measure discussed above. The overall gain for the Experimental groups (+12.46) is greater than that for the Control Group (+9.62), but a pooled variance estimate t-test for this comparison failed to reach significance (t[45] = .604, p = .549). There are obvious differences between the groups in Table 21, but these too do not reach a statistically significant level (F (3,45) = 1.56, p = .210). Once again there is the unexpected asymmetric profile of scores within the three Experimental groups, with I > II > III, that is, the Lo/Lo Ss make a higher percentage gain on this Efficiency measure than the other training groups. Despite the clear pattern of results here, none of the planned orthogonal contrasts produced significant differences between the groups. Post-hoc comparisons with Dunnett's t-test produced a significant result for the 12% difference in gain between groups I and III (t[4,45] = 1.98, p < .08).

The third method of calculating the Efficiency Ratio is the log transformation of the basic ratio score. The results of this are shown in Table 22, but they do not add anything beyond what the other measures of efficiency provide. The effect of the transformation is to greatly attenuate the data thus minimising even further the differences between the groups. None of the statistical tests of the data in Table 21 produced any significant results. The same order and pattern of results is repeated here as was demonstrated with the other efficiency measures. However, since this log ratio also produced a significant between-groups difference in the pre-test, thus negating the initial equivalence of the groups, it would appear that the gain-score ratio (discussed above) is still the most appropriate assessment of treatment effects for the Efficiency variable.

6) Ambiguity

The corollary of the Quality measure is, of course, the measure of amount of message ambiguity, expressed as the sum of the zero propositions across the three test items. Table 23 shows the post-test scores for this measure. All of the training groups have higher ambiguity scores than the Control group, but these group differences are not statistically significant ($F(3,41) = .617$, $p = .608$). Within the three Experimental groups, the Lo/Hi show the lowest amount of ambiguity.

More revealing results are provided by an analysis of the trends in change of performance from pre- to post-test. A gain score analysis of the data in Table 23 shows that the Training Groups as a whole show a positive change after training by decreasing the amount of ambiguous propositions in their message. This negative gain is greater than that shown by the Control Ss, who actually go in the opposite direction and increase their ambiguity score. These is also an apparent gain-score trend for Pairing effects within the three training groups, with I > II > III. The Lo/Lo Ss show a much greater decrease in ambiguity after training than the other two Experimental groups. However, despite these trends, univariate analysis of these training and pairing effects did not produce any statistically significant differences. For Experimental vs. Control, $F(1,45) = 0.211$, $p < .65$, and for I + II vs. Controls, $F(1,45) = 0.385$, $p < .54$.

7) Inefficiency

As was done with the measures of quality and efficiency, it is instructive to look at the degree of message inefficiency by adjusting the ambiguity scores for variations in message length. Even though the pre-test data showed no significant differences between the groups on the measures of ambiguity and quantity, there is enough variation in the pre-test scores for these two variables to warrant an inspection of Ambiguity/Quantity Ratio scores across the four groups.

Table 24 shows the post-test scores for the straightforward Ambiguity Ratio measure of Inefficiency calculated as the ratio of zero to total scores. All of the Experimental groups have higher Inefficiency scores than the Control group, but none of the group differences in Table 24 are statistically significant ($F(3,41) = .372$, $p = .774$). Within the training groups the Lo/Lo and Lo/Hi Ss have the lowest inefficiency scores. As with the measure of ambiguity (above), a gain-score analysis of the Table 24 data is more revealing about the positive effects of training on this variable. The Experimental groups as a whole reduce this inefficiency after training by an amount (-6.6%) greater than the change shown by the Control Ss (-3.4%). There is also the same asymmetry within the Training groups as reported before, showing the I > II > III trend for Pairing effects. The Lo/Los reduce their inefficiency after training by the greatest amount (-9.1%), the Lo/Mods are next (-7.9%), and the Lo/Hi's show the smallest drop in inefficiency (-2.9%).

Finally, the Inefficiency Ratio can be calculated via a log transformation of the basic zero/total ratio as was done for the Efficiency ratio discussed earlier. The results of this ratio transformation are shown in Table 25. As usual, the log transformation has greatly attenuated the scores and reduced the group differences, but the data still preserve the same pattern and order of results as was shown with the untransformed Inefficiency ratio measure above. The fact that the Control *Ss* had a low Quantity score and a low Ambiguity score to begin with leads to their having the lowest post-test ratio score in Table 25, but a gain-score analysis once again uncovers the positive effects of training and the relatively larger gains made by the Experimental *Ss* in reducing their propositional inefficiency.

8) Organisation

Having looked at the changes that take place at the propositional level *inside* the child's messages after communication training, we can now look at the organisational structure of the message as a whole. As before, the measure of message coherence is the supra-propositional index of Organisation. The post-training results for this variable are shown in Table 26. These data indicate that everybody improves their message organisation over time, with the Lo/Lo group outscoring all the other groups, and the Lo/Hi group having the lowest post-test scores. None of the post-test group differences in Table 26 turned out to be statistically significant ($F(3,41) = 1.62$, $p = .20$), even though the Experimental groups I and II scored higher after training than the Control group. The asymmetric pattern of results within the Experimental groups (I > II > III) again tends to attenuate Experimental vs. Control group differences. This same pattern, however, contributes to a significant difference *within* the Experimental groups. Group I (Lo/Lo) scores significantly higher than Group III (Lo/Hi) in message coherence after training (Dunnett's $t[4,41] = 2.35$, $p < .05$).

A gain-score analysis of the data in Table 26 once again produces the asymmetric pattern of results for the Experimental groups. Groups II (+4.66) and I (+4.46) again show a relative improvement from pre-test to post-test greater than that made by group III (+1.36). This pattern of results effectively washes out any significant differences when the average gain for all Experimental *Ss* ($\bar{x} = +3.49$) is compared with that made by the Control *Ss* (+3.46).

9) Composite Efficiency Measure

So far, we have looked at the different parts or segments of the child's First Message in terms of their quality, efficiency, organisation, etc. We can now turn, as we did in Part I, to look at the message as a whole, using our composite measure of overall "competence", which combines quality, quantity and organisational scores in one formula. What effect does communication training have on this composite measure? The results of the post-training evaluations are presented in Table 27.

These show a similar pattern to that encountered in the pre-test data. The range of efficiency scores remains approximately the same, from a minimum of 35% to a maximum of 95%. There is still a 10-point spread of scores across the four groups in Table 27, and the degree of variability is still high, especially among group III (Lo/Hi). All four groups have improved their efficiency over time and the average level of communicative competence for all 49 *Ss* is now 77%. None of the obvious group differences exhibited in Table 27 proved to be statistically significant ($F(3,41) = 1.02$, $p = .394$). Once again the Experimental groups, taken together, do not score significantly higher than the Control group, and this is again due to the asymmetric pattern of results within the Experimental groups (I > II > III), where the Lo/Lo *Ss* score highest and the Lo/Hi's score lowest.

Before considering a more appropriate gain-score analysis of Table 27, we can look briefly at the alternative formula for calculating composite Efficiency scores, viz., the log transform of the basic Composite score formula. The results for this measure are shown in Table 28. These data show the same relative positions and patterns as those in Table 27, yet the effect of the log transformation is to greatly foreshorten or reduce the results such that the more obvious group differences in post-test performance or amount of pre/post gain have been greatly attenuated. None of the results in Table 28 are significant, so it is more revealing to turn to an analysis of gain scores to ascertain where the real effects of training lie.

Table 29 shows the percentage of relative gain made by each of the four groups, and these results illustrate more dramatically and more significantly what is happening within and between the groups.

The differences between the groups in Table 29 are statistically significant ($F(3,45) = 2.91$, $p < .04$). However, the recurring asymmetric pattern of results is once again encountered within the three Training groups with I > II > III. Group III (Lo/Hi), who were the highest scorers in the pre-test, have made the smallest gain (+5.91%) of all four groups. As a result, the average gain for all Experimental Ss is greatly reduced and, while it is still slightly higher than the gain shown by the Control group, an orthogonal contrast of Experimentals vs. Controls proved to be non-significant. Other orthogonal contrasts made within the Experimental groups showed a significant set of ordered differences confirming the I > II > III pattern. For example, for Group II vs. III, $t(45) = 2.43$, $p < .02$, and for I vs. II + III, $t(45) = 1.74$, $p < .08$. Further exploratory, post hoc comparisons were made with Dunnett's t-test, and these revealed an additional significant difference between groups I and III, with $t(4,45) = 2.75$, $p < .02$. However, a comparison of the highest scorers, Experimental group I (Lo/Lo), with the Control group IV produced no significant results ($t[4,45] = .984$, $p < .10$).

10) Regression of Quality on Quantity

Finally we can look at one other measure of the effects of communication training on the propositional content of the subjects' verbal messages. This is the Predictor variable of propositional quality that is based on the regression analysis of the quality/quantity relationship. This relationship has been explored via our other ratio and composite measures of message efficiency and some significant trends and patterns have been detected in those analyses. Now we can see if the Predicted Quality scores cast any further light on these training effects.

Since the post-test correlations of quality with quantity are very similar with both $N = 49$ and $N = 97$, we took advantage of the added power of the larger sample and calculated the intercept and slope for the regression equation based on $N = 97$. For the post-test data, therefore, the Predicted Quality measure is:

$$\text{Sum Pos. Props} - (7.5 + (.4 \times \text{Sum Props}))$$

i.e., it is a child's total Quality score (number of positive propositions) corrected for the quality/quantity relationship (where Quantity is defined, as before, as the total number of propositions generated in the three tasks).

Table 30 shows the post-test scores for this predictor measure of message quality. All of the groups improved the quality of their propositions in relation to message length with the exception of group III (who were the highest scorers in the pre-test), who showed a very slight decline in performance.

In terms of absolute levels of performance the Experimental groups as a whole outscored the Control group, but none of the group differences in Table 30 reached a significant level in the analyses of variance. Within the Experimental groups we find the familiar pattern of I > II > III, with the Lo/Lo group outscoring the other two. A gain-score analysis of the pre-to-post changes shows even more marked differences between the training and non-training groups. The Experimental groups together have a much higher average gain in message quality (+1.25) than the Control Ss (+0.40). The asymmetric pattern of differential performance across the four groups is again repeated, with groups II and IV showing very little change, while the greatest gains are made by groups I and II.

11) Correlations between STL Measures

Finally, having looked at each of the STL variables individually, we can now examine the relations between these measures in a correlational analysis. A pooled within-group correlation matrix ($N = 49$) for the major post-test measures in the STL domain is shown in Table 48 . This was produced by the computer program as part of the multivariate analyses to be discussed in the next section.

We can examine briefly some of the highlights of these correlations from the post-training analyses. The reader should remember that these data are for all 49 Ss lumped together and that no distinctions are made between the Experimental groupings. Keeping this in mind, some general patterns are evident nonetheless. Comprehension of the basic prepositions is most strongly related to message Quality in both its raw ($r = .35$) and regression forms ($r = .41$), indicating, perhaps, the role of a precise vocabulary in formulating positive propositional messages. The measure of message Quality in its raw, uncorrected form, correlates highly with both message length and its own regression transformation. However, the nature of the quality/quantity relation changes when we inspect the log transformation of the measure of message quality. This ratio measure of Efficiency, corrected for differences in message length, has a negative correlation with Quantity ($r = -.49$), and a lower positive relation with the predictor measure of quality. This inverse relation between efficiency and quantity will be discussed further in a later section of the report dealing with factor analyses of the language domains.

There is a fairly strong correlation between the basic measure of Quality and the independent measure of message Organisation ($r = .55$). The nature of this correlation raises two important implications for our understanding of communicative abilities in the STL domain. First, it is obvious that the skill of formulating positive propositions is meaningfully related to the capacity for organising a variety of propositions into a coherent whole in a total message. The second implication, however, counterbalances the first. The size of the correlation indicates that each of these two measures, though related, still makes an important independent contribution to the total score pattern. That is, each measure is necessary, but neither one is sufficient for providing a well-rounded assessment of communicative competence: both together provide a balanced interpretation of message-construction ability.

This leads, in conclusion, to a truism (which is also applicable to some of the other relations demonstrated in Table 48) that the really competent child must master both the formulation of the parts of the message and the organisation of the parts into a coherent whole. The data suggest that these skills are not interchangeable, but they are measureable. The implication of these results for the future evaluation of communicative competence should be obvious.

12. Discussion of STL Results

There are a number of positive trends and some significant results for Training effects in the domain of competence that we have identified as spatio-temporal language use. One could be heartened by so many positive trends when one considers the small N that is involved, and it might be argued that more significant results could have been achieved with a larger sample. However, the major influence on these STL results is the unexpected pattern of Pairing effects within the three treatment groups (Lo/Lo, Lo/Mod, and Lo/Hi Ss), which essentially vitiated many of the planned analyses used to test these effects. The order of performance observed here (I> II> III) is contrary to the predictions from the a priori contrasts that were established before the experiment (Part I). At that time it was expected that all of the treatment groups would out-perform the control group after training, and the disappointing, almost atypical performance of the Lo/Hi group was not foreseen. In many respects, the original predictions for Pairing effects were the opposite of what was actually observed in the STL data, due to the combination of an unexpectedly strong performance of the Lo/Lo group and an unexpectedly weak performance of the Lo/Hi Ss. Despite their unexpected status, many of the Pairing results proved to be significant nonetheless, so there is an effect of Peer Modeling that has to be dealt with in the interpretation of the findings.

The most frequent victims of the lop-sided training-group results were the Experimental vs. Control comparisons that were used to test for treatment effects. On nearly all of the STL measures of communication ability groups I and II did better than the Control group, but the lower performance of Group III usually served to wash out this advantage when the average scores for the training groups were used in Experimental vs. Control comparisons. Before discussing possible interpretations of the Pairing results for the three training groups, we wish first to review some of the general findings for the STL evaluations as a whole.

If one thinks of competence in terms of fluency, then the post-test results would suggest that all 49 experimental Ss are quite competent indeed. The group as a whole produced an average number of about 23 propositions per child for the three-item STL task. The 36 trained Ss produced longer messages after training than the Control Ss, but the difference was not significant. Looking at the larger sample of Ss in the study (N = 97), we find that the Moderate competent and the High competent are all very "fluent" also if only length of message is considered. So it would appear that differences within and between levels of competence have to be sought in qualitative rather than quantitative variations in message content. Among the Low competent groups a high proportion (75%) of their propositions were of positive quality so these Ss would also appear to be very able when it comes to formulating communicative messages in this STL domain. There are obvious task constraints influencing message content in general here -- the Toy Village task is more structured, more closed-ended and more factual when compared, say, with the LL tasks. But the high Quality (or 'Truth') ratio in the childrens' messages here point to an important conclusion, viz., that memory is not an important factor in this type of task since so much of the information is accurately retained and reproduced.

Given the relatively high levels of competence demonstrated by all 49 Ss in the STL task, it is probably not too surprising that analyses using unadjusted post-test scores to measure differences in absolute levels of performance often failed to produce differences big enough or significant enough to differentiate between the four Experimental groups. Since the tests of the training effects were further compromised by the unpredicted pattern of Pairing effects already discussed, perhaps it is all the more surprising that any positive trends for the training groups were discovered at all. Yet the positive effects of training are visible in the STL data in a number of areas of competence: in the improvement of the propositional quality of message content in relation to message length, as attested to by the regression analysis; in the increase in message coherence and organisation; and in the overall balance of part-whole relations in message form and content, as delineated by the composite measure of competence. And these scoring trends and patterns were made even clearer, and proved significant in more instances, by the gain score analyses that were more appropriate to this STL data [this analysis strategy had implications for the evaluation strategies used in the Manova results to be discussed later]. This is particularly true for the significant Training ($p < .04$) and Pairing effects produced by the gain-score analyses of the composite measure of competence whose formula combines information on message quality, quantity and organisation. This variable would appear to be the one best measure of competence in this domain, but it owes its utility to the variety of fundamental measures on which it is based. The ability of these measures to illuminate different facets of the message-construction process points up the absolute necessity of evaluating communicative competence in this domain in a truly representative manner.

The correlational analyses of the various STL measures suggested that they all have a fair degree of independence and are measuring important and different aspects of message construction ability. When the data from these single measures are combined, they provide a richer, fuller picture of the pragmatic-linguistic complexity of children's oral communications. The present results provide some empirical justification for the argument that a complex process like CC (or any language process for that matter) cannot be measured by simplistic, single-item tests nor analysed via a single variable. The composite measure of STL competence which attempts to combine information from separate measures can be a powerful analytic tool. The findings from this part of the study should encourage multi-component evaluations and multi-measure analyses of competence.

Finally, how are we to explain the idiosyncratic pattern of Pairing results which showed the Low Ss to do best on the post-tests after being trained with other Low competent partners? Is the trend Lo/Lo > Lo/Mod > Lo/Hi real or is it an experimental artifact? Well, in one respect, it seems largely due to the atypical performance of the Lo/Hi Ss, who, in fact, showed a comparatively higher degree of variance than the other groups in their scores on such measures as Efficiency, Organisation and Composite Efficienty. It seems fair to say that, if the Lo/Hi group had consistently performed at the same level as the Lo/Mod and Lo/Hi Ss, then most of the Training effects would have been significant for the Experimental groups. So perhaps we have simply run up against the perennial problem associated with this type of experimental design, viz., that a treatment group (the Lo/Hi's) could perform in the post-training evaluations in a way that is less reliable and less truly representative of what they showed themselves capable of doing during the actual training sessions. However, to say that the Lo/Hi Ss simply had an "off day" does not answer all the questions. When the performance of the larger sample (N = 97) was checked for the ratio and the composite measures of efficiency, the results ran counter to the usual 'High > Mod > Low' trends. They showed the Low competent groups to be performing higher than the Moderate groups who had better scores than the High groups, and the pattern 'Low > Mod > High' was statistically significant. Are we to conclude, therefore, that, contrary to the original predictions there can actually be a negative effect of peer-modeling on the learning of this particular type of English usage which has a natural S-V-O sequence? On the face of it the results certainly suggest that a Low group's success on the STL tests has an inverse relation with the level of ability of the training partners in the three experimental groupings. It is true that in the original proposal the hypothesis concerning the direction of influence of High competent peer models on Low competent trainees was left open-ended, since it seemed possible that a match-up of extremes like this might be counterproductive, viz., that the Low Ss would be out of their depth, so to speak, with the speech input from their sophisticated peers going over their heads during the communication training sessions.

But this kind of speculation would have to be stretched uncomfortably far to account for all the Pairing effects encountered in the post-test data. We would prefer to stay closer to home and look for methodological or procedural factors that may have negatively influenced the post-training test performance of the aberrant Lo/Hi group of Ss. Since all of the experimental groups were proven to be initially equivalent before training, and since all three received the same training program, could there be anything in the post-training situation that might have influenced the responses of the Lo/Hi Ss? A re-analysis of the test-partner pairings in the post-test situation strongly suggested to us that familiarity of test-partner may have had a determining influence in the level of test-performance of the Lo/Hi group. We will explore these familiarity factors in more detail in the next section in an attempt to discover a more parsimonious interpretation of the unique Pairing effects in the STL data.

13. Dyadic Testing and Familiarity Effects

A unique feature of the language evaluations in the present study was that children were tested, in pairs, with other children. This type of dyadic testing creates the possibility that a target child's test performance might be influenced not only by antecedent treatments (e.g., the experimental manipulation of level of competence of training partner), but also by situational, testing conditions (such as the degree of familiarity or strangeness of the child's test-partner). We have already recounted (Section III. I. 1. ii) the elaborate precautions that were taken during the post-training evaluations to control familiarity effects in various ways (a test-partner could not come from the same classroom, or be a training partner, and so on.) Table 1-b showed test-pairings broken down by competency status, but there are degrees of test-partner familiarity mixed in with all these dyads which have to be looked at in more detail.

A basic issue here is generalisation of learning: were any of the STL test-partner pairings beneficial/detrimental to any experimental Ss such that they did/did not allow these trainee Ss to show the maximum benefits of training during the post-training evaluations? The more anonymous the test-partnerships the more the STL battery becomes a test of generalisation of practice effects to novel conditions. Now all four Experimental Low groups met a baseline condition for equality of testing conditions to the extent that no S could ever be tested with a partner from his own training group. This meant, therefore, that the post-test conditions ensured a minimum degree of generalisation of trained competence from all Experimental Ss, since everyone had to communicate to a "strange" (i.e., non-training) partner. However, the degree of "strangeness" of test-partner apparently varied widely across groups.

The dyadic testing data of Table 1-b were reanalysed and broken down by Competency and EXPGROUP status for both partners. These re-analyses are shown in Table 1-c, and this shows that, at the most general level, the procedural attempts to control test-partner pairings were moderately successful, as far as creating objective, quasi-anonymous dyads was concerned. Out of the 39 test-dyads involving Low Ss, 10 (26%) involved pairs of Low Ss, 15 (38%) were pairings of Low and Moderate Ss, and 14 (36%) paired Low and Hi Ss. So far, so good for the generalisation issue -- Low Ss had to interact more often with non-Low test-partners in the testing situation. Further success was achieved in controlling for familiarity by balancing the proportion of Experimental and Control group Ss with whom the Low Ss were paired. Out of the 49 Low Ss, 25 were paired with Experimental Ss and 24 with Control Ss (Table 1-c).

The problem of familiarity effects becomes clearer, however, when we look at the structure of the test-partner pairing within each of the four Experimental groupings. The patterns in the data of Table 1-c lead to a number of conclusions.

(1) First, there appear to be three kinds of Familiarity Effects that have to be considered:

(A) Personal Familiarity: having a partner you know (classmate) or have worked with in communication task situations (a training-group partner). This kind of familiarity was minimised to a large extent by our experimental controls.

(B) Ability-level Familiarity: Is a S paired with someone at his own level of competence? [another sub-category here might be: being paired with someone at the level of the subject's peer models in training].

(C) Task-knowledge Familiarity: here the effect is weighted in favor of having a partner who is an Experimental rather than a Control S. That is, someone who has been in the training program experiencing tasks similar to the test problems, whether or not he or she is at the same ability-level as you.

Looked at in this way, maximal test-partner familiarity would be an "A+, B+, C+" condition, i.e., being tested with a partner you know, who is at the same competence level as yourself and who has been in the training program. Minimal familiarity would mean A-, B-, C-, i.e., being paired with a stranger, of a different competence level, who has not been in the training program.

(2) The results of the partnership breakdowns in Table 1-c indicate that the latter condition of minimal familiarity was the more frequent fate of group III (Lo/Hi). The data suggest that the previously observed pattern of Pairing effects

(I>II>III) may be largely due to the fact that the dyadic testing structure tended to maximise familiarity effects for group IV, minimise them for III, and enhance them positively for group I. This, in effect, stacked the cards against the Lo/Hi Ss, while providing test circumstances for the Lo/Lo and Lo/Control Ss which would facilitate their performance.

The details of Familiarity Effects B and C for each of the four groups is as follows. For group I (Lo/Lo), while the Experimental vs. Control split is fairly evenly balanced (46% vs. 54%), the largest proportion of test-pairings here is Low with Low (46%), thus favoring Familiarity Effect 'B'. For group II (Lo/Mod), the balancing of both Familiarity Effects, 'B' and 'C' is perfect. The E vs. C balance of partners is 50/50, and there is an equal proportion of exposure to partners at each of the three COMPSTAT levels -- one-third Low, one-third Mod, and one-third High (Table 1-c).

With groups III (Lo/Hi) and IV (Control) we see the extremes of the Familiarity effects, positive for IV and negative for III. Of the three training groups, the Ss in III had the least amount of test-pairings with partners who were at the competence level of their own training models: group III had only 27% such, compared to 33% for II and 46% for I. Group IV likewise had a majority of its pairings with other Low Ss (46%), and less with other levels of COMPSTAT, thus getting the benefits of Familiarity Effect 'B'. Familiarity Effect 'C' was also at work on the balance of Experimental vs. Control group partners for group III Ss. The latter's test-pairings were made even more anonymous and objective here, with eight of the eleven Lo/Hi Ss (or 73%) being tested with partners from the Control groups (Lo or Mod or Hi) who had a minimum of task knowledge or sophistication. The opposite was true for the Lo/Control Ss in group IV where ten out of the thirteen (77%) were partnered with Experimental subjects who had some task sophistication to bring to the testing.

(3) Overall, therefore, compared with the other Experimental groups, the Lo/Hi Ss had the least advantageous testing experience. They had less opportunity to work with partners at the same ability level (Low), or with partners at the same level as their training peers (High), or with partners who had some experience or sophistication in the type of test problem. Group I (the best post-test performers), on the other hand, were favored by having more opportunity to interact with test partners at their own level of competency. Similarly, the group IV Ss, who had the least amount of task knowledge, were paired with other Ss who had the most amount of such knowledge. What would have happened to the test performance of group III if their familiarity effects had been more equitably balanced? To answer this we can look to group II (Lo/Mod) as a model, since they had an ideally balanced set of test-partner pairings. On virtually all of the STL measures the gain score analyses in Tables 18-13 showed group II outperforming group IV. Were the Lo/Hi Ss to perform as well as these prototypical Lo/Mods, the average scores and gains for the three training groups together would have been consistently higher than those of the Control group, and many more significant treatment differences would have been found.

B. Logical Language (LL)

1. Summary Overview

When we take a preliminary overview of the basic post-training evaluation results in the LL domain, the profile of test scores is quite different from that encountered in the STL domain of competence. As a matter of fact, the pattern of results provided in Table 31 is virtually the inverse of that previously illustrated in Table 16 for STL data.

Leaving aside treatment differences for the moment, when we look at the pooled data for the 49 Ss in Table 31 we can see, for instance, that the LL tasks appear to be a more difficult battery of tasks for the Ss than the STL items. Here the level of mastery is still fairly low when we compare the proportion of positive propositions to ambiguous ones in all of the First Messages — it is approximately 33% to 66%, though this still represents an improvement from the pre-test where the Positive/Ambiguous ratio was approximately 25% to 75%. So there has been a change in Message Quality and Efficiency over time and this appears to have been accomplished without any change in Message Length or Fluency.

As far as Quantity is concerned, the total number of propositions (1647) has changed very little from the pre-test evaluations. Each S appears to have added about one extra proposition to his repertoire as a whole, with the average number of propositions per child for all three tasks being 33.6. This represents an average of 4.2 propositions per item for each of the eight test items, which is half of the item average for the STL domain.

Not only is there little change in Message Length over time, but the Quantity measures, both within and across the three LL tasks, show remarkable consistency across time. Each item still contributes between 10% and 16% of the Total Quantity Score, while the three sub-tasks each contribute ±33% to its own task total. For measuring Message Fluency or Quantity at least, the LL battery shows good psychometric stability and good test-retest consistency.

What does Table 31 tell us about qualitative, substantive aspects of message content? If the number of propositions in the First Message remains relatively stable, wherein do the post-test changes lie? Table 31 would indicate that they seem to lie in a moderate increase in message quality and in a decrease in message ambiguity. The ratio of positive to ambiguous propositions now changes to 34% vs. 66% respectively, so that one-third of all messages are now unique propositions. Using the percentage of Positive Propositions per Task as an index of TASK DIFFICULTY, the three tasks are now ordered, from easiest to hardest -- Cartoon Boxes, Production Board, and Picture Boxes. The order of ITEM DIFFICULTY by the same index undergoes some changes and is now #3 (Hardest), 5, 2, 6, 4, 8, 1, 7 (Easiest). An item x task analysis of Table 31 reveals that the most dramatic increase in message quality is to be found in the three-fold increase in positive propositions in the hardest task, Picture Boxes. [If other analyses confirm that the effect of training is to increase the quality of performance in the harder items, as was shown with the cognitive data earlier, this would provide a strong justification of the linguistic benefits of oral communication training.]

The most striking improvements in performance across tasks are on the harder items, viz., Picture Boxes items #3 (the hardest), where the proportion of positive propositions goes up from 1% to 10%; #4, where the change in message quality is +23%; and #5, where it goes up +13%. The increase in the proportion of positive propositions for the Picture Boxes task as a whole, for all 49 Ss, is +14%, which represents the largest post-test change in message quality for any of the three LL sub-tasks.

In sum, therefore, Table 31 allows us to say in a very general way, that there is an overall increase in message quality in the LL domain. It now remains to be seen how these post-training results break down in terms of Experimental and Competency status and across the various measures of message quality, ambiguity, efficiency and organisation.

One final comment can be made, en passant, about the data of Table 31. At the broadest level of analysis it reveals that, when one samples the language of 8-year old children across a wide enough range of communication tasks of the LL variety, their functional, pragmatic speech is shown to contain a good deal of ambiguity.

N.B.: Test Order Effects

For all of the major, summary variables to be discussed below, there were no statistically significant results for Test Position effects (Form A vs. Form B). Consequently, unless otherwise noted, equivalency of Test Forms will be assumed for the purposes of analysis and discussion of the post-training results. The discussion to follow will, therefore, concentrate on within- and between-group experimental effects.

2) Quantity

For the LL domain, as with STL, the measure of Message Fluency or Length is the total number of propositions utilised across the eight items of the three tasks, regardless of content. Table 32 shows the post-test results for this measure.

Variability among scores decreases in the post-test and the range of fluency scores (17 - 52) decreases at the upper end of the scale. As Table 32 indicates, Quantity or amount of propositions per message remains fairly stable from pre to post, for each group. The Control group still has the smallest average number of propositions (30.15) and Experimental group I (Lo/Lo) still has the highest number (36.92). In absolute terms, the three Experimental groups have longer messages, on the average, than the Control Ss, after training. However, none of these differences between post-test scores in Table 32 was statistically significant.

In relative terms, the most notable gain from pre- to post-test was made by group III (Lo/Hi Ss), who increased their message length by almost three propositions. Overall gains, however, for the three Experimental groups taken together are approximately equal to the average gain for the Control group.

3) Message Discriminability

This discrimination index is a summary measure of the highest level of propositional quality achieved in a child's First Message. As such it is a very specialised measure of the categorical quality of a child's message.

Table 33 shows the post-evaluation scores on this measure for the four groups. The counterbalancing of test order during the post-training evaluations has apparently successfully cancelled out the Position effects ($B > A$) which were found to be significant in the pre-test. A test of Position effects for the post-test data was non-significant ($F = 0.156$, $p = .695$ (1,41)).

There are, however, significant main and group effects for differences in discriminability scores in Table 33. After communication training, the scores of all three Experimental groups are consistently higher than those of the Control group, and these group differences are significant ($F = 3.43$, $p < .026$ (3,41)). The highest scores after training are achieved by groups III (Lo/Hi) and II (Lo/Mod).

The Experimental groups as a whole also show greater gains on the average (+5.06) than the Control group (+4.46) from pre- to post-test. Group III (Lo/Hi *Ss*) outstrip all other groups in the extent of their increase (+6.64) in the discrimination quality of their messages over time. In both absolute and relative terms, therefore, it is these Low *Ss* who are paired with High competent peer models who show the strongest positive response to training.

4) Message Quality

Another measure of message quality in the LL domain is the number or frequency of unique propositions in a child's messages across all of the eight items in the LL tasks. The post-training results for this summary measure are shown in Table 34.

There is no significant Position effect ($F = .238$, $p = .629$ (1,41)) here as there was in the pre-test results, indicating the successful effect of counterbalancing test order in the post-training evaluation. All *Ss* improve their scores by three or four points on this measure, but it is the three Experimental groups, both separately and together, who score markedly higher in the post-test than the Control group. The highest average number of unique propositions (12.09) is attained by the Lo/Hi *Ss* of group III. Despite the obvious trends in Table 33, however, none of these post-test group differences proved to be statistically significant ($F = 1.58$, $p = .209$ (3,41)).

It is the Experimental groups as a whole who also show the greater relative gain (+4.08) from pre- to post-test when compared with the Control group (+2.69). Both absolutely and relatively, therefore, the Experimental *Ss* outperform the Control *Ss* on this measure of message quality.

5) Efficiency

The measure of message quality discussed above did not take into account individual differences in length of message. When we correct Quality scores for differences in number of propositions on each item for each child, we effectively create a ratio measure of message Efficiency. The post-training results for this measure are shown in Table 35.

The trends in these data are somewhat similar to those revealed in the Quality results above. All Ss improve their message efficiency over time, while group variability remains stable. The post-training scores in Table 35 for the Experimental groups are uniformly higher than those for the Control group, but none of these group differences are significant (F = .148, p = .930 (3,41).

The relative gain made in message efficiency from pre- to post-test is slightly larger for the Control group (+1.01) than for the average gain of the Experimental groups taken together (+0.83). Among the Experimental groups, only group I (Lo/Lo) shows a gain (+1.02) comparable to that of the Control Ss.

6) Inefficiency

The converse of efficiency is, of course, inefficiency or ambiguity, i.e., how many non-unique propositions the child produces in his first message on each task. As before, the measure of inefficiency is the ratio of ambiguous propositions to total number of propositions in each first message, summed across tasks.

The results for this Inefficiency measure are shown in Table 36. There is a general trend for all Ss to reduce their ambiguity over time. At the post-test level, all the Experimental groups score slightly lower than the Control group, but none of these post-test group differences in Table 36 are significant (F = .077, p = .972 (3,41).

From pre-test to post-test, group I (Lo/Lo) and group IV (Controls) show the greatest amounts of negative gain, or reduction in ambiguity, with -1.02 and -1.06 respectively. The average negative gain for the three Experimental groups taken together, however, is smaller (-0.77) than that for the Control group (-1.06).

In general, the results for this measure show no clear-cut trend in favour of the Experimental groups.

7) Organisation

Previous measures, like Quality and Discriminability, have looked at the incidence and frequency of unique and non-unique propositions in the child's first message. The measure of message Organisation now looks at the location of unique propositions within the first message across tasks, to determine if training has been effective in changing the position of such propositions inside the message.

Table 37 presents the results for this measure of message organisation. Across all Ss, it is obvious that the location of a unique proposition has moved up closer to the front of the message, on the average, thus rendering the latter more informative in the early states of the communication. There are differences between the four groups, however, in the extent of this shift, and these differences proved to be significant.

All three of the Experimental groups have lower post-test scores than the Control group, meaning that the messages of the training subjects were better organised in terms of the more forward location of their unique propositions. These between-group differences are significant (F = 2.55, $p < .06$ (3,41)). Within the Experimental groups, it is group II (Lo/Mod) who give the strongest organisational performance, with their first unique proposition being located, on average, at approximately the fourth proposition into the message. This can be compared with the virtual sixth proposition location of such utterances for the Control group.

A similar pattern of results is observed when one looks at the relative changes in performance over time. From pre-test to post-test, the Experimental groups taken as a whole do a little better, on the average, that the Control Ss. Within the Experimental groupings the strongest shifts occur in groups II (Lo/Mod) and I (Lo/Lo).

In general, there is no clear-cut ordering of Pairing effects within the three Experimental groups. The most noteworthy feature of the results in this regard is the consistently strong performance of group II (Lo/Mod Ss), both in terms of absolute level of post-training performance and of relative gains made over time.

Overall, this measure of message organisation provides positive results for the effects of training on the informational quality of the children's communications. As a summary measure it provides an average profile of quality of message organisation across all test items and tasks. It was noted above that the results for this summary measure of organisation indicated a shift in location of unique propositions towards the front of the message, thus ensuring added clarity and non-redundancy in the early part of the child's communication. By comparison, an inspection of the sub-task post-training data for this variable confirm the general trends reported for the summary measure, but the location of unique propositions in the post-test and the extent of their shift appears to be a function of the difficulty level of each task.

In each task there is a positive trend, which favors the Experimental groups, towards lowered location scores after training. In the easiest task, Production Board, there was, on the average, a shift in location of unique propositions from the middle to the first quarter of the message. In Cartoon Boxes, a like shift is from the middle to the first half of the message, on average. Perhaps the most interesting data come from the post-training results for the hardest task, Picture Boxes. This was the task for which, at the pre-test, the strongest case for initial equivalence of groups could be made, since the spread of scores across the four groups was only a half a point or less. At the pre-test level, unique propositions were either non-existent in the messages of most of the subjects in this task, or they occurred way at the ends of these communications. At the post-test, on the average, the location of such unique propositions had moved towards the middle of the first message.

It is also worth commenting here that the most positive effects of training were visible on this Picture Boxes sub-task — the hardest in the LL domain, and one which should have provided the toughest test of communication learning. Both absolutely (in level of post-test score) and relatively (in amount of pre/post gain) the Experimental groups as a whole outscored the Control Ss. Furthermore, within the Experimental groups, a more predictable pattern of Pairing effects was observed in this sub-task, with the Ss in groups II and III, who were paired with more competent peer models performing better than group I (Lo/Lo Ss).

8) Regression of Quality on Quantity

Finally, we can look at the predictor relationship between Quantity (total number of propositions across all first messages in all three tasks) and Quality (total number of positive propositions) in the children's communications. As we have noted earlier, message length in the LL tasks was much more constant and less variable across time than it was in the STL domain. [The correlation between Quantity and Quality in the LL domain is still very similar at the post-test ($r = .629$) to that of the pre-test ($r = .658$).] As such, it could be argued that there is less necessity to measure the effects of Quantity on Quality in the LL domain. Nevertheless the use of this regression analysis for the LL data can be informative for at least two reasons. First, it enables us to make comparisons, between the two language domains, with the similar predictor variable used in the STL analyses. Secondly, within the LL domain, it complements the earlier analysis of the Quantity/Quality relationship provided by the ratio measure of message Efficiency by looking at the same data in another way.

As it turns out, the regression analysis provides a clearer picture of the effects of training on message quality than that provided by the previous measure of Efficiency. The regression equation for the post-test data is:

Predicted Quality = (Total # Positive Props.) - [3.3 + (0.2 x Total # Props.)]

The data for this predictor measure of message quality are shown in Table 38.

The significant Position effects encountered in the pre-test results appear to have been successfully controlled by the counterbalancing of Test-Order in the post-training evaluations and there are no significant test position effects for the data in Table 38 ($F = 1.16$, $p = .288$ (1,41). Since there were no significant Group effects or interactions in the pre-test data, initial equivalence of groups was assumed, and we can look at the absolute level of performance in the post-test as an index of training effects for the Experimentals vs. Controls comparisons.

As Table 38 indicates, the trends in the post-training evaluations favor the three Experimental groups. All Ss improve their Quality scores, but only the Experimental Ss now show a positive relation of Quality to Quantity. For the Control group this relationship remains negative (-0.02). Despite the obviously higher level of performance by the three Experimental groups, none of the post-test group differences in Table 38 proved to be significant ($F = 1.17$, $p < .333$, (3,41)). There also appear to be an ordered set of pairing effects within the Experimental groups, with $III > II > I$, i.e., with the Lo/Hi Ss turning in the highest performance.

The same pattern of scoring trends is repeated when we look at the relative amount of score gain made by the groups over time. From pre-to post-training, the Experimental groups, taken as a whole, make almost twice the amount of gain (+2.86) in level of message quality that the Control group does (+1.50). This time, however, the Pairing effects within the Experimental groups are not so predictable, with groups I and II making relatively greater gains than group III.

Overall, this regression analysis of message quality in relation to fluency shows a clearer pattern of scoring trends in favor of the three training groups who outperform the Control group in both absolute and relative terms.

9. Correlations between LL measures

The pooled, within-group correlation matrix for the major LL variables is shown in Table 48, for N = 49. The magnitude of the intercorrelations in this domain is generally higher than it was between the STL measures. The basic measure of LL Quality correlates fairly well with Quantity in the post-test (r = .586), but appears to be virtually interchangeable with the predictor measure of message quality with which it correlates r = .917. This is not the case, however, with the summary ratio measure of Efficiency (which is actually Quality corrected for differences in message length) which shows an inverse relationship with message length (r = -.197). There appears to be good internal consistency among the various measures of message quality, with the measure of Efficiency correlating strongly with propositional quality, in both its raw form (r = .546) and its regression corrected version (r = .752).

Table 48 also provides encouraging evidence of the complementary nature of many of these measures which appear to be tapping overlapping characteristics of message-construction ability in this LL domain. There are, for instance, strong relations between message structure (Organisation) and message content (the various measures of Quality and Efficiency) where the average r = .7, and interesting parallels can be drawn with similar comparisons in the STL data. The Organisation variable would appear to play a pivotal role in the LL analyses since it correlates well with not only the pragmatic-linguistic measures of Quality but also with the Discriminability Index (r = .75) which has an added perceptual-cognitive dimension. In this regard it is interesting to note that this Discrimination measure, in contrast to its strong relations with message content, and message structure, appears to have relatively little to do with message length (r = .256).

The overall pattern of the intercorrelations in Table 48 suggest that the more discriminating variables in these LL analyses are those dealing with quality of message content and structure. This finding, coupled with the strength of the correlation between many of the variables, provides further evidence of the reliability and utility of the LL test battery for measuring functional language abilities in this domain.

10. Discussion of the LL Results

As a separate domain, the LL results invite some interesting comparisons with the levels of performance in the STL domain. It should be remembered, however, that the LL analyses are based on the use of summary measures which themselves are based on many more observations from many more test items (8) than were used in STL (3). It seems very likely that the greater strength of the intercorrelations of LL variables as compared to those between STL measures is at least partially due to this larger data base.

At the grossest level of analysis, looking at all variables for all Ss, some general trends are apparent in the LL data. Message fluency (Quantity) remains fairly stable over time and the changes that occur appear to do so in qualitative aspects of message content and organisation.

These changes are positive on the whole, but the overall profile of competence for this kind of language use is still relatively weak, and about two-thirds of the message content produced by the Experimental Ss is, on the average, ambiguous. Of course content was measured by very conservative criteria of "quality", since only unique, non-redundant propositions were scored positive. A less rigorous scoring system might have produced a different profile of competence here but at least the present system has the minor merit of avoiding the "grey areas" of measurement by scoring as positive only those utterances that are clearly at the top of the scale of quality. A conservative scoring system can run the risk of underestimating ability, especially if it is used on a short test with few items. But the LL battery has enough range and variety of tasks to minimise this risk, and the speaker has many opportunities to display his competence on test items with varying degrees of difficulty. So it can be claimed with some confidence that the test profile provided by the LL measures is a good representation of the range of competence in the sample. It is also noteworthy, that an ordered hierarchy of difficulty is still preserved across tasks and items in the post-test data, thus providing one indication of good test-retest consistency. Another indication of strength in the LL test battery is its sensitivity to training effects, as there are a number of indications in the data that the level of performance of the training groups improved most strikingly on the harder test items (e.g., on the Picture Boxes task), which is what one would predict if truly progressive new learning had taken place. To these psychometric virtues of the LL test battery we can also add the evidence of strong intercorrelations between the major LL analytic variables. All in all, therefore, it would appear that the LL test battery has the potential to be a useful diagnostic tool in this domain.

Looking now at the data separately for the four Experimental groupings, the positive trends for training effects seem even clearer in this domain than they did in the STL data. The Pairing effects, while not always clear-cut, do not have the systematically deviant pattern they showed in the STL results and, as a result, differences between Experimental and Control groups in average level of performance are allowed to show through more clearly on most of the variables. These differences were generally also large enough and consistent enough that they could be demonstrated quite simply through the use of unadjusted post-test scores, with less need of the gain-score analyses which proved to be more appropriate for the complicated STL data. [Gain score analysis might still be a useful strategy for unravelling some more treatment effects for one or two of the LL measures and this will be considered in the MANOVA analyses to be discussed later.] Not all of the obvious group differences observed reached statistical levels of significance, but we might cite again the small sample size as a possible factor in this. Sample size notwithstanding, there were significant results for training on two important variables, the measures of categorical quality (Discrimination, $p < .03$) and of organization of first messages ($p < .07$). Once again, as in the STL domain, these LL results point up the importance of measuring both message content and message form. The intrinsic nature of the relation between what is said and how it is said is further underscored by the results of the correlation analyses which showed that the ability to construct a discriminating proposition is highly and significantly related to the ability to place that proposition early in the message where it will have greater communicative value for an active listener. Future research with these variables might profitably experiment with combining these two measures in some composite measure of LL competence along the lines of the successful formula used with the STL measures. The profile of competence provided by the LL measures likewise reinforces the argument made for the STL results above that the complexities of functional language use in this area are best rendered and illumined through a variety of measures which together provide stronger results than single measures alone.

The measures of propositional Quality and message Efficiency all showed positive trends for Training effects and this is especially apparent on the regression measure of Efficiency (which is predicted Quality with the effects of the Quality/Quantity relationship removed). The raw score measure of message Quality (which is uncorrected for variations in message length) also correlates well with both Quantity and the predictor measure of Efficiency, and provides important information on its own, independently of any consideration of quantitative variations in messages. There are grounds, therefore, for considering it as a useful variable to include in the confirmatory Manova analyses to be discussed later.

Finally, what about Pairing effects? These were not evaluated statistically at this level of analysis of the LL variables. This was because the exploratory univariate analyses were organised to focus on P6sition effects as well as EXPGROUP effects, since there had been some significant Test Order effects among the pretest measures. Pairing effects for the LL variables will, however, be evaluated in the Multivariate analyses to be discussed later. The descriptive data in Tables 32-38 nevertheless show some patterns and trends that allow us to make some general comments about the order of Pairing effects for LL. There is less of a consistent pattern as there was with the STL pairing data, but at least there is not the aberrant performance of the Lo/Hi group that there was in the STL domain. The unadjusted post-test scores give the clearest picture of systematic differences in levels of performance among the four Experimental groups. On all seven of the LL variables, group IV is the lowest performer, while among the three Training groups the highest scoring on nearly all of the measures is usually produced by the Lo/Mods and Lo/Hi's. The latter group, interestingly enough, give the best performance on three important measures — Discrimination, Quality, and the predictor measure of message efficiency. It is on this latter measure of predicted quality that we find the only perfect instance of a hierarchically ordered set of Pairing effects among the four groupings, with the levels of performance running III > II > I > IV. The pattern of modeling effects is less clearly ordered in a consistent manner across the rest of the LL measures except in the grossest sense already referred to, that groups I, II, and III always out-perform the Control group IV.

Gain score analyses proved to be of little value in uncovering any systematic patterns of Pairing effects. About the only result that can be noted with any confidence here is that fact that of the two extreme groups in terms of levels of peer-modeling — Lo/Hi and Lo/Control — group III showed greater positive change after training than group IV on four of the LL measures. There are no other clear trends within the three Training groups, and we will have to consider next whether or not there were any Familiarity Effects influencing (positively or negatively) the performance of the four Experimental groups.

11. Dyadic Testing and Familiarity Effects

The basic question here is, as before, whether or not some accident of the test-partner pairings during the post-training evaluations unduly facilitated or hindered the extent of competence displayed by the Ss in the different Experimental groupings. Table 1-d provides some answers to this question. The results of the test-partner pairings are skewed less dramatically in the LL domain, but there is still evidence of some familiarity effects which might influence the strength of the between-groups effects that were tested in the LL data. Although the effects are less extreme than they appeared in the STL data, the overall assessment of the partnership data in Table 1-d suggests that group III (Lo/Hi) still seems to be least favored and the Control Ss (group IV) most favored in so far as enhancement of responsiveness by familiarity effects is concerned.

At the most general level of analysis, the balance of the dyadic test pairings was fairly even across the three ability levels for all 49 Experimental Ss. Table 1-d shows that there were 40 test dyads involving Low competent Ss. Within these, 18 (37%)Low were paired with other Low competent partners, 19 (39%) Lows were paired with Moderate competent partners, and 12 (24%) with High competent Ss. Thus 31 out of the 40 dyads (roughly 75%) tested Low Ss with non-Low partners, so one could argue perhaps that in one sense at least the general test-partner structure guaranteed a minimisation of familarity effects and a maximisation of objectivity and generalisation for the Low Ss. In addition, their test partners were well balanced between 20 (41%) Experimental Ss, and 29 (59%) Control or Spare Ss.

However, within each of the four Experimental groups the pattern of partnership was neither so well-balanced nor clear-cut. Looking at Familiarity Effects 'B' (ability level) and 'C' (task sophistication), the data in Table 1-d indicate that there is still some enhancement of effect 'B' for groups I, II, and IV. The Lo/Lo's had more opportunity to interact with partners at the same ability level, while the Lo/Mods had the most opportunity to be tested with partners at the same level of competence (Moderate) as their peer training models. However, these positive type-B effects for these two groups were cancelled out by their negative type-C effects since 3/4 of all of their partners were naieve, non-experimental Ss. Group III, the Lo/Hi Ss who had the deck stacked against them, so to speak, in the STL test-pairings, do somewhat better in the LL testing. Type 'C' effects are more equitably balanced this time across Experimental, Control, and Spare partners, but type 'B' effects were more inequitably distributed for these Ss than for any of the other groups. Only 9% of their test-partnerships gave group III the opportunity to interact with partners who were at the same level of competence as their peer-training models, compared to 38.5% for I, and 42% for II.

Group IV (Control Ss), on the other hand, has a skewed distribution of partnerships similar to the one encountered in the STL data, and one that still tends to maximise both familiarity effects 'B' and "C" for this group. For example, the Ss of this group were most often partnered with Ss at the same ability level as themselves (46% of the time), while 10 (77%) of their 13 partners were the more task-sophisticated Experimental Ss.

In conclusion, therefore, the structure of the dyadic testing in the LL domain seemed to accord a more conservative treatment to the three Training groups (especially to group III) in so far as benefitting from familiarity effects is concerned. One could argue, for instance, that the nature of the Training groups' testing situation made greater demands on their abilities to generalise their newly learned competence to "new" partners. The Control Ss, on the other hand, were made to "look good", as it were, by the facilitating nature of their post-test partnerships. It seems entirely plausible that these kinds of familiarity factors could have an influence on the LL post-training evaluation results for both training and modeling effects. Yet the very fact of the existence of such factors, and of the direction of their influence, makes the positive trends visible in the LL training data all the more significant.

C. Summary Discussion and Implications of the Language Domain Results

The foregoing discussions of the results of the post-training analyses for the two separate domains of Spatio-Temporal and Logical Language usage highlight a number of similarities and differences between the domains, and we will attempt to briefly summarise and synthesise the main issues here. Some of the findings have immediate, practical implications for the kinds of analysis strategies to be adopted in the confirmatory, multivariate analyses to be discussed in the next section of the report. However, we will also attempt to touch on more conceptual issues which have longer term implications for theory, research, and educational practice in this area of functional language development.

(1) Regarding similarities and differences between the domains, one could point to individual measures whose results seem to shade the profile of competence differently in each domain. In the less structured LL tasks, for example, one of the most significant variables is the Discriminability Index which measures a child's success in imposing organisation and structure on his verbal message by differentiating and categorising the most distinctive features of the target picture. In the highly structured STL tasks on the other hand, it is an aggregate measure of propositional quality, quantity and organisation together, which is the most significant differentiator. However, at an even more general level, one of the more striking differences between the two domains is in the overall levels of competence manifested in each, when we compare the relative proportions of propositional quality across all messages. While virtually 3/4 of the Experimental groups' utterances in the STL tasks are of positive quality, almost the inverse is true in the LL domain. We have already mentioned the different task constraints in the two domains which might contribute to the different manifestations of communicative competence in the two domains. But we should also point out that these differences are also partly due to the differing communicative demands imposed by each task. Different language usages are required by each task, and the children obviously dealt better with the pragmatic language requirements of the STL tasks than they did with those of the more open-ended LL tasks where they had to impose their own structure and organisation on their message forms. The correlation between the major cognitive measure and the LL organisational variable suggests that there may be a cognitive-organisational component underlying success on the LL tasks, and this would account for the better performance in this domain of High competent Ss and the Lo/Hi Experimental Ss.

There is one immediate implication of the differences in competence in the two major language domains displayed in this study. Many previous studies (e.g., Glucksberg, Kraus and Higgins, 1975) have chosen to measure children's oral communication competencies using only referential communication tasks from the more difficult LL domain. Our findings now suggest that these one-sided studies have grossly underestimated the child's general communicative competence, and that reliable generalisations about levels of communicative ability have to be based on samplings of more than one domain of functional language use.

The fact that our two major domains elicit different kinds of speech also has developmental implications. Sinclair (1975) has suggested that STL is more akin to the child's natural language which (in English) has a similar, left-to-right, subject-verb-object sequence. As such it is learned earlier than language usage of the LL variety which contains more idiosyncratic forms and appears to conform less naturally to the child's familiar structural-linguistic forms. Logical language usage appears to be harder to acquire and to be more dependent on cognitive organisation and level of intellectual functioning than STL. If this interpretation is correct, acquisition of mastery of the two domains would appear to be differentially influenced by environmental and educational factors, and this is a topic to which we shall return later.

(2) The exploratory analyses of the Quality/Quantity relations in the two domains have direct implications for the confirmatory, multivariate analyses to be presented in the next section of the report. In neither domain did message length (Quantity) seem to be a major differentiating variable among the four Experimental groups, but we still experimented with different ways of looking at message Quality (number of positive propositions), either in its pure form or transformed into various measures of Efficiency which all attempted, in one way or another, to correct Quality scores for variations in message length. These supposedly more economical and parsimonious ratio measures of Efficiency were not necessarily more informative about treatment differences. In some cases (e.g., the log-transformed ratio of STL Efficiency) we seem to have over-corrected the quality scores with the result that real group differences were greatly attenuated. This was also true in LL where the adjustments for quantity variations seemed redundant on tasks like Picture Boxes where there were relatively few positive propositions to begin with. Given the positive linear relationship between Quantity and Quality in both domains, and also given the lack of significant group differences in message length, there are grounds, therefore, for using the basic untransformed measure of Quality by itself, in both the exploratory and confirmatory analyses, for looking at absolute levels of message quality across the groups. Additional support for this decision is provided by the high correlations in both domains between the raw measures of propositional Quality and the measures of Predicted Quality which have the effect of the quality-quantity relation removed. By contrast, the major Efficiency variables in both domains have a negative relation with message length and in order to interpret the meaning of this relationship we will have to consider these Efficiency measures as separate variables in the multivariate analyses to follow.

(3) Another result which also directly affects the strategies chosen for the confirmatory, multivariate analyses is, of course, the nature and order of Pairing effects which were observed in the two domains. These effects directly influenced the size and significance of many of the Experimental-versus-Control group comparisons which were made to evaluate Training effects. This was particularly true in the STL results and different interpretations can be made of the Pairing effects that were shown in both domains.

Part of the problem lies in the original hypotheses and experimental contrasts which were constructed on the premise that all of the Training groups would outscore the Control group on all of the variables in all of the tasks. These initial predictions did not foresee that the study would uncover two separate language domains which might require different and separate sets of predictions. They certainly did not foresee the atypical performance of the Lo/Hi Ss which rendered many of the planned a priori comparisons ineffective or counterproductive in the STL analyses. By use of different kinds of alternative strategies and experimental contrasts a number of positive and significant trends for training effects were uncovered in the exploratory analyses, and similar tactics will have to be adopted in the confirmatory, multivariate analyses for assessing the combined effects of individual variables within and across domains.

The Pairing effects in the LL data were not as clear-cut as those observed in the STL results, but they still showed trends in a direction different from that seen in the STL data. If one were to take the trends in the two domains at face value, one would have to arrive at two opposite conclusions: that STL training effects can be maximised for a Low S only if he is paired with another Low competent partner, but that pairing him with a High or Moderate competent training partner produces better results in the LL domain.

We have tried to argue, on the other hand, that many of the pairing results could be due to a factor unique to this kind of dyadic testing, namely, to the familiarity effects that derive from pairing a subject with certain kinds of testing partners in the post-training evaluations. We have described in detail what these familiarity effects might be, but it remains to be determined how these facilitating (or hindering) effects are actually transmitted during the testing session. We feel that this is a subject for further research, but we might offer some speculations about how familiarity effects might positively or negatively enhance an experimental subject's test performance. In the positive instance, a sophisticated test partner might actually be engaged in a sort of "modeling-during-testing" which could facilitate the other subject's responsiveness. In the negative case, where the least familiar test conditions exist, a subject's responsiveness may be inhibited due to a lack of the "shared pre-suppositions" between partners which Bates (1976) claims are the pre-cursors of effective communication.

Whatever the reason, this newly discovered Familiarity effect appears to be a variable that is unique to this kind of training model and experimental design. As such it is a methodological variable that will have to be controlled in future research. One can manipulate it in one direction to maximise test responsiveness, or in the other direction to create post-training testing situations which provide the strongest challenge to generalisation of learning.

(4) Another striking feature of the training results in both domains was the important roles played by variables dealing with both the form and the content of communicative messages. We have argued accordingly that a profile of communicative competence must talk about message structure as well as message quality, and the importance of these part-whole relations in the construction of oral messages has implications for both measurement and theory in pragmatic language studies.

The measurement issue is one of representativeness of assessment and analysis techniques. The complexities of functional language cannot be captured by single-measure analyses. A wide-ranging profile of competence comes from assessing a representative cross-section of abilities with a variety of tasks and from analysing different facets of performance with a variety of measurement techniques. In this regard we feel that the results from the post-training evaluations in both language domains reinforce the findings from the pre-test Canonical Correlation and Discriminant Function analyses that it is groups of measures which give a more powerful interpretation of "competence" than any single item or variable.

Our observations of the importance of part-whole relations in message formulation also raise theoretical considerations regarding the conceptual analysis of an oral communication at different levels. The propositional analysis scheme that was developed in this study provides a segmental analysis of the units of a message. It now seems necessary to look at the child's oral message as a whole, that is, in a supra-segmental way. We would like to propose, for example, that, phenomenologically, the message as a whole is greater than the sum of its parts. Conceived of in this way, not only the speaker but also the listener must be responding to the message at two levels. How could this be evaluated?

Sinclair (1980) has recently suggested an approach to discourse analysis that could be fruitfully applied to the interpretation of children's dialogues. Sinclair's approach goes beyond strictly linguistic analyses of discourse content and applies more psychological, Piagetian categories to the child's utterances. The child's speech may be grammatically immature but Sinclair's segmental analysis of the child's message to another provides fascinating examples of the existence and manifestation of three kinds of knowledge at work. Parts of the child's discourse may show, in differing degrees and at differing levels of development : first, Mundane Knowledge (knowledge of the world, objects, things); next, Interpersonal Knowledge (social relations); then Operational Knowledge (the laws, rules, relations governing things). This is an extremely helpful way of looking at message content and the development of functional language ability. Yet it still does not deal with the message-as-a-whole, which we feel is the best expression of the child's communicative intent.

There are meta-segmental or meta-linguistic issues here which require a pragmatic analysis that is quite different from a structural analysis, whether linguistic or psychological. The total message, for example, can reflect a level of cognitive, perspective-taking ability in a way that no single segment could. Hopefully, future research might be able to apply Sinclair's scheme to a supra-segmental analysis of a child's message, asking how the message-as-a-whole reflects mundane, interpersonal, or operational knowledge. Conceived of in this way, an oral message is not a purely linguistic phenomenon, but is also a cognitive and social event, and must be analysed as such. There are three issues that might be worthy of further study: 1) How does the child-speaker develop and show awareness of supra-segmental factors? 2) How does the child-listener develop and show comprehension of supra-segmental factors? 3) How can this knowledge be tapped experimentally, through the use of appropriate tasks?

(5) Finally, what are the educational implications of the findings from the language domain? Can the functional language competencies identified as necessary in the LL and STL domains be integrated into a language arts curriculum?

There are some direct recommendations that can be made about the teaching and evaluation of basic skills in the classroom. For one thing, we would like to see the present study have an influence on the nature, variety and extent of oral language practice in the classroom. In addition, if the minimal competency movement in the United States is intent on devising tests of Speaking and Listening as well as of Reading and Writing, these tests must take account of the organisational components which we have found to be an important aspect of oral language competence. Multiple-choice, true-false, short-answer tests simply will not tap the organisational skills that are a necessary component of oral message construction. If, as some current educators are recommending, we need more essay questions in writing tests, then we will need more oral discourse questions in the assessment of speaking and listening.

Finally, we hope that the findings from our language domain will encourage a closer integration of teaching practices for speaking and writing. In advocating a language experience approach to elementary English education the Bullock Report (1975) claimed that the spoken word must mediate the written. The data from the LL domain suggested that there may be a cognitive component to oral language organisation. We would predict, therefore, that oral training in Logical Language use would have a positive effect on LL writing, and vice versa by fostering this organisational process. The discipline and practice required by LL writing should move students away from more concatenated, STL-type sentences to produce richer and more complex descriptions.

III. Results of Primary Measures: Confirmatory Analyses

Introduction

So far we have looked at the perceptual, cognitive and linguistic variables individually, attempting to identify which measures show trends and patterns, and applying univariate analyses to assess the significance of response differences on the separate measures. This has helped us to select major variables for follow-up analyses, and to preview the kinds of experimental contrasts and analysis strategies that will be most appropriate for the next stage of confirmatory analysis.

It is now time to move from the single to the multi-variable level of analysis in order to test how well or how poorly our multi-component model of communicative competence hangs together. We will use Correlation, Factor Analyses and Multivariate analyses (with some supplementary univariate analyses) to determine the significance of the observed scoring patterns within and across domains, and what they can tell us about the effects of communication training.

III.1. Post-Training Intercorrelations

We will first look at the interrelations between domains using the pooled within-groups correlation matrix (N = 49) produced by the MANOVA program (Table 48). This matrix corresponds closely to the post-test correlational data for the total sample (N = 97), so it will be used here both for a profile of the post-training interrelations and for comparison with the pretest matrices. The data in Table 48 show the 15 major variables to be studied in the confirmatory analyses -- two Perceptual, one Cognitive, six STL and six LL. [This post-test matrix contains two extra measures, the predicted Quality measures for STL and LL.]

Since we have previously looked at the correlations within domains, what does Table 48 tell us about relations between the four major domains of competence? There are at least two kinds of information that one can look for in the post-training correlations: one is the change in inter-domain relations over time and what the resultant correlational patterns might tell us about the nature and extent of possible training effects; the other is using the size of the correlations between the language domains, on the one hand, and the perceptual and cognitive domains on the other, as a clue to the "loading" of the latter on the former, in order to ascertain which STL and LL message construction abilities are tapping into more perceptual skills and which into more cognitive ones. It should be remembered in all of this discussion, of course, that the data in Table lumps experimental and control Ss together, and some of the changes in correlation patterns probably reflect some confounding of effects (though the pooled within-groups algorithm has attempted to deal with this to some extent).

An overview of the results generally indicates a mixture of changes in the Perceptual and Cognitive spheres, while in the language areas there is a general weakening of relations both within and between the STL and LL domains, with some exceptions which will be discussed later. With regard to the questions raised just above concerning relations between the language-training and non-language training domains, the Perceptual component appears to have strengthened its relations with the STL domain more than with LL, while the reverse is true of the Cognitive component. Within the Percpetual domain, it's the measure of Perceptual Accuracy which shows the more interesting changes in its interrelations. The perceptual measure of Speed of visual selection still has very low correlations with most of the other variables, though it does show a modest increase in correlation with its companion measure of Accuracy (r = .23). The latter perceptual measure shows its most substantial change in its relation with STL Efficiency, which it increases to r = .40 (p<.01). It also shows a fair degree of relation with the new regression measure of STL Quality (r = .28), while it improves to a lesser extent its correlation with accuracy of Prepositional Usage. This latter measure shows a similar mild increase in its relation with the Cognitive domain, and what is probably involved in both these relations is improved precision of vocabulary having an effect, covertly or overtly, on task behavior.

The most striking change in Cognition - STL relations is the inexplicable drop in the predicted relation of STL Organisation x Cognition from pre- (r = .23) to post-test (r = -.03). Other relations between cognition and STL are generally weak and, by contrast to this, the Cognitive component of competence increases its correlations with the LL variables all across the board. Most notable among these are its relation with the Discriminability Index (r = .33, p <.03) and message Organisation (r = .39, p<.01).

In this whole chain of Perception-Cognition-Language relations running through the post-training matrix, the relation between Perceptual Accuracy and the Cognitive measure is of particular interest because it appears to create an important "bridging" effect between the two language domains. There is a substantial change in this Perception x Cognition relation over time, so that now r = .34 (p<.02). This is consonant with the results of the previous sub-analyses which showed that increased cognitive performance is associated with fewer choices, slower times and increased accuracy in the perceptual domain. What makes this correlation particularly fascinating is that, in the light of the previous discussion, it raises various possibilities of language training effects being mediated indirectly through Cognition to Perception, or vice-versa. Take, for instance, the Perception x LL relation which is weaker than might be expected. The LL Discriminability Index is substantively based on accuracy of visual target analysis where distinctive, differentiating perceptual features are encoded into a verbal message form. Intuitively, at least, one would predict a strong correlation between this measure and the Perceptual component after training, yet the data for N = 49 in Table 48 does not show this. What we do see, however, is that LL Discrimination and Perceptual Accuracy both correlate equally strongly with cognition. Does this mean that the Cognitive component is mediating the effects of LL training to the Perceptual Domain? It is always risky to pre-judge the direction of influence via correlation analyses and, since this study's model of communicative competence claims that its component processes operate in parallel, one would expect the path of interrelations here to be a two-way street. Depending on which direction of influence one chooses to emphasise, one could claim either a) that LL training indirectly influences perception via changes in decentering and perspective-taking behavior; or b) that it is the highly analytic LL domain which benefits from an increase in perceptual accuracy when the cognitive component mediates in some way between the perceptual analysis of a target picture and the linguistic description of it. There is probably some truth in both these accounts, but the second interpretation would certainly be favored by the increased association of Cognition with LL Organisation as well as with Discrimination. It makes sense to think that the cognitive component would be involved not only in the selection of distinctive propositions about target features but also in their judicious location in the structure of the message.

If these are the possible trends of training influence between LL and Perception/Cognition, what then is happening between the latter domains and STL? It is probably dangerous to assume that the same kinds of effects are being transmitted or mediated in the same way or for the same reason here as they were in the case of LL. While the bridging that is taking place between the language and non-language domains may appear to share a superficial similarity for both STL and LL, the underlying explanations that are needed to explain the mediation of effects may be quite different for each language domain. This is especially true for the post-training language data since the correlational patterns show STL and LL to be moving towards increasing independence of each other. The Perception x STL relations, for instance, may not involve so much the kind of perceptual analysis we argued for in the case of LL as they do perceptual recall. This would certainly provide a more plausible explanation of the increased association between STL Efficiency and Perceptual Accuracy ($r = .40$, $p < .01$), given that the STL tasks are probably less dependent on analysis of features than they are on organisation of memory for action sequences. The organisational components of this recall process might then be mediated back from Cognition via the Perceptual component to improve STL message Quality by reducing the number of inefficient descriptions that a partner has to listen to.

There is obviously some speculation involved in these interpretations of the language/non-language inter-domain relations, but the fact remains that STL relates directly to Perception but indirectly to Cognition, while the opposite is true for LL.

Finally, there are the relations between the two language domains to be considered separately from Perception and Cognition. What do the correlations tell us about the influence of the two kinds of functional language usage on each other? Table 48 indicates that the STL and LL measures still show stronger relations within their own domains than they do with each other. However, the general pattern of weakening of inter-domain relations has its exceptions. The STL Prepositional Usage measure has dramatically increased its correlations with LL Efficiency and Discriminability with whom it now associates as strongly as it does with basic message Quality in its own domain ($r = .35$, $p < .02$, in all three cases). It also now correlates moderately well with the new predictor measure of Quality in STL ($r = .41$). The common element in all of these interrelationships is most likely to be increased precision of vocabulary affecting propositional quality of messages in both domains. Another variable which runs counter to the common trend is STL Quantity which surprisingly increases its relations over time with LL Quantity, Organisation and Discrimination. The most obvious reason we can think of for this is that language training has a generally positive (though limited) influence on message fluency in both domains. Within its own domain, STL Quantity increases its correlation with basic message Quality ($r = .70$) but it also strengthens its negative relation with STL Efficiency ($r = -.49$). [This latter result is corroborated by data from another correlational analysis for $N = 49$ which showed that STL Quantity also correlates -.46 with the untransformed ratio measure of STL Quality, virtually unchanged from pre to post.] Quality and Efficiency also continue to be negatively related in LL, while LL Quality has a lower but still significant relation with Quantity. In both domains the untransformed measures of message Quality have become virtually interchangeable with the new regression measures of predicted Quality (in STL, $r = .86$; in LL, $r = .92$).

The transformed measure of STL Efficiency has weakened all of its relations over time with variables in both language domains, and it now correlates best with Predicted Quality in its own domain (r = .37). On the other hand, Efficiency in the LL domain has increased its already strong relations with Organisation (r = .71) and Quality (r = .54) while demonstrating a strong correlation with the new regression measure of Quality (r = .75). LL Quantity, Organisation and Discrimination all tend to weaken both their intra- and inter-domain relations over time, but the relations between these latter two variables, and between them and the other LL variables, still continues to be the strongest of any domain.

Finally, what about the status of message structure, and its relation to message content? We have already noted that LL Organisation correlates quite well with other LL variables as well as with Cognition. In STL, message Organisation still correlates more strongly with its fellow variables, expecially with Quality of message content (r = .55, $p < .01$), once again drawing attention to the importance of the form-content relation that we have noted in previous results. The relation between Organisation and message length remains constant (and moderate) from pre to post in STL, suggesting perhaps that the structuring of a message requires a minimal number of propositions to work on. The relations of STL Organisation with the LL domain are still low, and its correlation with its counterpart, LL Organisation, has fallen to almost nil in the post-test. This is probably another indication of the separation of the two domains under the influence of training.

SUMMARY DISCUSSION

Since the experimental and control groups are lumped together in the data of Table 48 , there are obvious limits on the interpretations one can make about the visibility of training effects in the correlation data. However, some of the weakening of the relations between domains (especially in the language areas) would seem to be due to the effects of training separating variables along certain dimensions, given what we already know about significant group differences from previous univariate analyses. We will briefly summarise here what we consider to be some of the more important highlights in the changes in correlation patterns discussed above.

The relative independence of the two language domains in many respects continues to be demonstrated, and we will attempt to confirm this trend in the Factor Analyses of the language domains to be reported in the next section. However, from a psychometric point of view, the pattern of relations within and between the two language functions is what one would hope for, since it is obviously preferable to have the intra-domain relations be higher than the intercorrelations between domains for the sake of content validity.

The patterns of correlations in the post-training period indicate that different things are happening in each language domain and these will probably require differing explanations of the effects of training for each domain. The relations of STL and LL to the domains of Perception and Cognition, for instance, seem to have different implications for the independence or interdependence of language and thought in these two areas of pragmatic language functions. We have already mentioned the different kinds of mediating roles that might exist between LL, Cognition and Perception, and between STL, Perception and Cognition — roles which ultimately have implications for the nature of the relation between STL and LL themselves.

The Cognition-LL relation is of particular interest, for both intuitive and theoretical reasons. We noted above how cognition provided a "bridge" between Perception and LL, and how the cognitive domain appeared to have stronger relations in general with LL than with STL. There seems to us to be at least two issues of message organisation and content involved in these relations. We noted, first, that the kind of feature differentiation required by the LL target pictures has a high visual-perceptual content. Next, we observed in previous analyses that the unstructured aspects of the LL tasks require a greater application of the child's cognitive-organisational and perspective-taking abilities. There is a suggestion in all of this that some of the pragmatic-linguistic effects of training of LL may be more indirect and may be less "linguistic" than they are "pragmatic." That is, what is being changed is the child's ability both to select what is important to include in a message and to place it in a meaningful location inside that message. This requirement may be less relevant in the STL tasks where the nature and sequence of important information has been largely pre-determined by the experimenter.[1] This would also partly explain the lower association between Cognition and STL.

Another way of expressing these language-cognition relations is to suggest that the communication training affects the child's "knowledge" (in Sinclair's sense) of what a 'message qua message' is supposed to be like, as opposed to his knowledge of structural linguistic features. This points up the useful distinction between knowing what to say (linguistic)and knowing how to say it (pragmatic).

Not all of the relations between STL and LL are necessarily indirect or mediated via Perception and Cognition. The correlation data indicates that the Discriminability variable in LL has a possibly pivotal role in transmitting training effects directly from one language domain to the other. The Discrimination measure of categorical quality of LL propositions correlates as well with basic propositional Quality in STL (r = .32) as it does with Cognition. We interpret this to mean that increased precision in LL usage raises the chances that propositional content will also become increasingly precise in STL. The claim that this reflects a genuine effect of training is reinforced by the results of the earlier univariate analyses which showed the LL Discriminability Index to be one of the variables wherein performance differences between Training and Control Ss were statistically significant. In addition, the kinds of language-cognition relations we have been discussing here are also implied in the results of those earlier analyses. They showed that the highest LL Discrimination scores were turned in by Group III, whose peer models were High competent Ss who had the largest endowment of cognitive perspective-taking ability measured in this study.

Finally, what also seems to be common and advantageous to both language domains is possession of a useful vocabulary, as measured by the STL Prepositional Usage measure. This has apparently little to do with fluency, but it does contribute to accuracy of propositional content in both STL and LL. In view of this inter-domain relationship, STL Prepositional Usage should be a good candidate for inclusion in the confirmatory MANOVA analyses.

---

[1] This immediately suggests a topic for further research -- would there be more involvement of Cognition in STL if the children were left to make up their own action sequences as they went along?

III. 2. Factor Analysis

Introduction

Additional investigation of the uniqueness or interdependence of the two language domains was made by means of Factor Analysis. Confirmatory analyses were made, using Principal Components and various orthogonal and oblique rotations, to verify the structure and pattern of the relationship between the two domains as suggested by the correlational analyses just discussed above and the results of the pretest Factor Analyses. As in the pretest, the factor analysis of the post-training data takes advantage of the larger data base provided by all 97 of the experimental and control Ss. The correlation matrix for the post-test Factor Analyses for N = 97 is shown in Table 50. (A comparison of this matrix with the correlation data for N = 49 for the same variables in Table 48 , shows the two sets of data to be comparable in most respects). As with other analyses, the strategy for choosing variables for the factor analyses (as well as for the multivariate analyses to be reported later) again requires the use of summary measures which group the item and task measures in order to improve the power and sensitivity of these variables as domain measures.

As was explained in Part I of this report, the early, exploratory factor analyses were largely used to boil down and simplify the large amount of pre-test data generated by the cognitive, perceptual and linguistic measures. En route to identifying patterns of relations among these variables, these early factor analyses operated on groups of measures drawn from as many as 76 separate item and task scores. These analyses usually threw out about 14 to 16 factors which generally supported the multi-component model of communicative competence. These pretest results consistently showed separate factors for the Perceptual, Cognitive, STL and LL domains. The number of factors identified and analysed often depended on how many separate scores had been entered from the LL test battery (which, of all the domains, produced the largest and most varied assortment of measures).

Later pretest analyses concentrated on a reduced set of summary variables from the important STL and LL domains which were to be the major targets of the subsequent confirmatory analyses of training effects. As Table 49 indicated, four variables from each language domain (representing Quality, Quantity and Organisation) were entered into the final pretest factor analysis which produced four factors: a Quality factor for each domain, and separate Quantity factors for STL and LL.

Post-test Exploratory Analyses

Before turning to the post-training replication of these STL/LL factor analyses, we will look briefly at some exploratory analyses with the post-training data from all the domains. These were done only for purposes of comparison with the similar pretest analyses which had served to establish at the outset that there were separate factors for each of the four domains of competence predicted by the model. The data of these post-test exploratory factor analyses will not be presented here, and we will restrict the discussion to summarising the results. Using the post-test data for N = 97, 43 variables were entered into a Principal Components analysis. In addition to variables from the Perceptual, Cognitive and STL domains, this included 30 different item and task measures from the LL domain. Varimax rotation produced 14 factors which included, in addition to Cognitive, Perceptual and STL factors, eight separate factors for the LL tasks. This analysis simply confirmed the existence of separate factors for each domain, but it was not pursued any further because of the decision to work with summary measures to simplify the factor structure within the language domains.

Post-test Confirmatory Analyses

The existence of discrete Perceptual and Cognitive components, each with separate and unique functions, continued to be confirmed both in the exploratory analyses above and in the post-training correlational analyses discussed in the previous section. The latter analyses also indicated a growing independence between the two language domains, so a factor analysis was undertaken to confirm this pattern of relationships among the language variables deemed most important for the assessment of training effects. Eight variables were entered into this analysis—Prep. Usage, Efficiency, Organisation and Quantity from STL, and Discriminability, Organisation, Efficiency and Quantity in LL. Table 50 shows the resultant factor structure. Analysis by principal components and varimax rotation again produced a good cluster solution which essentially replicates the pre-test findings. Four factors were again identified.

The First Factor is a LL Quality factor loading on Discrimination, Organisation and Efficiency. The Second Factor is an STL Quantity/Efficiency factor loading highly positive on Efficiency and highly negative on Quantity. The Third Factor is an STL Quality factor loading on Prep. Usage and Organisation. The Fourth Factor is an LL Quantity/Efficiency factor, loading highly positive on Quantity and negative on Efficiency. We thus appear to have for each language domain two matching dimensions, one being overall Quality of message and the other an inverse relation of Fluency to Efficiency. The one difference from the pre-test is that the Second and Third factors have changed positions. Further oblique rotations confirmed the factor pattern, leading to the general interpretation that the two domains do indeed retain their independence over time.

The Quantity/Efficiency inversion is also much stronger in both STL and LL in the post-test. This result may seem at odds with the correlational data which has shown on numerous occasions that the raw untransformed measures of propositional Quality correlates positively and strongly with their respective Quantity counterparts in both domains (in STL r = .7, and in LL r = .6). Perhaps these two variables are so much alike that, whatever they are measuring, the effect is washed out when they are both combined in a ratio score. There is enough contrast in these results to argue that raw Quality and Quantity ought to be examined separately from Efficiency in the subsequent multivariate analyses. It should also be remembered that even a correlation of .6 between Quantity and Quality still leaves 40% of the variance unaccounted for, so it is probably not too surprising that we get the four factors that we do in the principal components analysis.

In conclusion, the post-test factor analyses have continued to show separate factors for the four domains, and have produced a clear-cut pattern of independence between the two language domains along dimensions which are substantively similar. This latter result confirms trends indicated in prior correlational analyses. It now remains to be seen in the subsequent multivariate analyses whether communication training has had direct or indirect effects within and between these independent language domains.

III. 3. Confirmatory Multivariate Analyses

i) Introduction

In the previous sections of this part of the report we have presented results from one-way ANOVAS, Correlations and Factor Analyses. These have all been steps along the way to this final stage of confirmatory analysis of the Training and Pairing effects in this study. We are now about to embark on a final series of Multivariate analyses of the post-training data in order, first of all, to evaluate the general theoretical model that underlies this whole study, secondly, to test the specific experimental hypotheses and assumptions with which we began this study, and thirdly, to attempt to confirm or disprove some of the additional, unpredicted trends and patterns of performance that were shaken out of the data along the way during our exploratory analyses. Before outlining the scope and sequence of the multivariate analyses which will be addressed to these major issues, we should briefly review and recapitulate the highlights of our findings up to this point.

Our theoretical model of communicative competence posited two non-language domains, Perception and Cognition, which were explicitly involved in the measurement and evaluation of competence in the first phase of this study. However, the content areas of these two domains (as measured in this study) were not explicitly involved in the communication training program and, as such, the post-test perceptual and cognitive measures might be thought to reflect some degree of generalisation of learning effects from the program. In this regard, the Perceptual results were not too revealing and the univariate analyses produced ambiguous or inconclusive findings for both training and pairing effects. Much more clear-cut, and significant results were produced for Cognition where predictions were confirmed for both the influence of training (Experimental > Control) and peer-modelling of competence (groups III > II > I). Further Correlational analyses also showed the cognitive component to be more strongly related to LL than STL, and it appeared to mediate between the Perceptual component and some aspects of LL functioning. Perception, on the other hand, showed more direct relations with STL.

The univariate analyses of the language domain results produced different patterns of responding within STL and LL. In STL, unpredicted patterns of performance by the Lo/Lo and Lo/Hi Experimental Ss affected most of the contrasts of Training effects and produced an unpredicted order of Pairing effects within the Experimental groupings. Nevertheless, some significant effects of training and pairing were still found among the raw post-test scores on some of the measures. We attempted to attribute the unique and unpredicted patterns of results in the STL domain to procedural or methodological factors, citing the influence of familiarity effects on test performance. As a consequence, the real meaning of Training and Pairing effects in STL will have to be evaluated further in the confirmatory analyses by looking more closely at gain scores and using different kinds of experimental contrasts.

Training effects seemed to be more clearly and predictably established by the univariate analyses of the LL results, particularly on important measures of categorical quality and organisation of messages. Pairing effects were not always clear-cut, but the Lo/Hi Ss often turned in the highest levels of performance on a number of post-training measures. The more predictable and clearer pattern of significant training effects in LL was demonstrated by analysis of unadjusted posttest scores after initial equivalence of experimental groups had been established at the pretest level. Accordingly, it will be less necessary to consider other types of scoring strategies (e.g., gain score analysis) in the confirmatory, multivariate analyses of the LL domain below.

Correlational analyses of the post-training scores from the two language domains for all experimental Ss (N = 49) showed some general trends towards the weakening of relations between STL and LL over time. Nevertheless, there could also be discerned in the intercorrelations a positive (although limited) influence of training common to both domains in such areas as message fluency, increased precision of vocabulary and some aspects of message quality. Principal components analysis further confirmed the relative independence of the two language domains, and highlighted the inverted nature of the relation between message length and certain measures of message efficiency. While basic propositional quality and quantity still remain strongly and positively related in both domains, both factor analyses and regression analyses indicated that lengthening of messages was not the most direct way of increasing message quality and this is particularly true of the LL domain where the children appeared to be less "competent" in the efficiency-ratio sense to begin with. The results of the analyses of the quantity/quality relations in the language domains have obvious implications for the selection, grouping and ordering of variables in the various multivariate analyses below.

ii) Overview of the Analyses

To provide the reader with an overview of the analyses below and offer a guiding rationale for the selection of the various MANOVAS which will operate on the four domains, separately and jointly, we will summarise briefly the issues to be discussed below.

a) Selection of Variables

The fifteen variables to be studied have already been discussed in the previous correlational analyses (Table 48). All of these variables are summary measures having as much independence as possible in formulation and calculation from other variables in the same domain. There are two Perceptual measures of Speed of target detection and overall Accuracy of performance, along with one summary Cognitive measure. All three of these variables were chosen on empirical and/or substantive grounds to be the best representatives of their respective domains. There are six STL and six LL measures representing comparable aspects of message structure, content and length (or fluency) in each domain. Included in the STL set is the measure of accuracy of prepositional usage which was shown to have important lexical relations with indices of message quality in both language domains. Both domains also contain the different measures of propositional quality, either in its basic, summary, uncorrected form, or in an efficiency-ratio form corrected for individual differences in message length at the item or task level. Also included are the new post-training regression measures of Quality corrected for Quantity. Each of these differing indices of quality of propositional content will appear in separate, non-redundant groupings of STL and LL variables in the analyses below in order to answer questions about their different roles as indicators of treatment effects in the study.

b) Selecting the Model of Communicative Competence

The preceding exploratory analyses have raised new and unforeseen questions about the nature and structure of the model of communicative competence (MCC) that underlies this study. The selection of analyses and contrasts below should provide some partial answers at least to some of these questions. For example, if we are still dealing with the sort of well-balanced tri-partite model by means of which we conceptualised the study at the outset, then we should expect the MANOVA results below to produce lots of multivariate joint effects showing that training is maintaining and strengthening inter-domain relations all across the board. However, we have now found that the language domain is less monolithic than originally conceived and now contains at least two separate language functions. This means that the model now contains four components rather than three, which raises further questions about the nature and strength of inter-domain relations. We should now, perhaps, look for fewer joint effects between language domains, and expect to see more effects among the sub-components of Quality, Quantity and Efficiency within each of the domains separately.

And how does the newly-dichotomised area of language function interact with the other components of Perception and Cognition in the model? If we have pragmatic differentiation within the language domain, perhaps we should also expect that the interaction between the language and non-language domains will be further differentiated also. For instance, the data to date have suggested that Perception and Cognition relate in more or less direct ways to STL and LL usage, raising the possibility, heretofore unconsidered, that the effects of training will be mediated in different ways at different levels of interaction within the model. We should expect, therefore, that MANOVA'S which include or omit Perceptual and Cognitive variables in different kinds of groupings of STL and LL variables should produce different types and degrees of effects both within and between domains.

Whatever the outcomes of these analyses, the results will most certainly lead to a revision of our conception of the dynamics of our model of communicative competence.

c) Analysis strategies and the choice of experimental contrasts

The original study proposal called for the use of orthogonal Helmert contrasts to test Training and Pairing effects, and we will continue to use these as the contrasts of choice, together with program-generated SIMPLE contrasts (Finn, 1974) of experimental vs. control group differences wherever appropriate.

For four experimental groups, the HELMERT CONTRASTS are:

CONTRAST A: All Experimental groups vs. Control group.
(I + II + III vs. IV)

CONTRAST B: Lo/Lo *Ss* vs. Lo/Mod + Lo/Hi *Ss*.
(I vs. II + III)

CONTRAST C: Lo/Mod *Ss* vs. Lo/Hi *Ss*
(II vs. III)

Contrast 'A' represents the test of treatment or Training effects between trainees and controls, while Contrasts B and C represent tests of Pairing effects within the three experimental groups.

The SIMPLE Contrasts are pair-wise comparisons of each experimental group with the control group, i.e., Contrast D = I vs. IV, Contrast E = II vs. IV, and Contrast F = III vs. IV.

The Helmert contrasts will be used for group comparisons of raw, unadjusted post-test scores where initial equivalence in performance has been established across the four groups. Both the Helmert and Simple contrasts will also be used to evaluate, in certain instances, interesting gain-score changes in performance from pre- to post-test. This will be done to confirm additional important training or pairing effects that were suggested by the prior univariate analyses.

The STL data will require different treatment in the analyses. We have already seen that the asymmetric pattern of responding produced by the three Training groups renders comparisons like Helmert Contrast 'A' ineffective and inappropriate as a test of real training effects, since the exceptionally low performance of Group III pulls down the Experimental Groups' average in such a comparison. This has the effect of masking real training effects that may be occuring in Groups I and II. To offset this we will test treatment effects for STL with a set of ARBITRARY Contrasts which will incorporate a more meaningful comparison of experimental with control groups. In addition, those experimental contrasts will operate on gain scores for the STL measures since the exploratory analyses indicated that measurement of relative changes in performance over time ought to give a fairer and clearer picture of the influence of training on the experimental groups.

d) Supplementary Univariate Analyses

We will also supplement the multivariate analyses by discussing univariate results for a number of variables. These one-way Anovas will either be those produced by the MANOVA program or they will be supplementary analyses carried out to answer specific questions about scoring trends for particular variables (e.g., the Cognitive measure), whether this means differences in absolute levels of post-test performance or in relative score gains over time.

In the case of the Anova produced by the MANOVA program, we will be looking at selected univariate results whether there is a joint effect or not. It is not usually legitimate to look at significant univariate results when the multivariate effect is negative, but we have some valid reasons for doing so. In the previous exploratory univariate analyses most of the tests were done by two-way Anovas, which looked at Test Position effects as well as group effects. Since all of the Position effects were non-significant in those analyses, we now have an interest in confirming predictions made at that time regarding potentially significant group effects for particular variables. The one-way Anovas produced in the MANOVAS provide exactly the between-groups tests we need for these with the appropriate degrees of freedom. All comments on these univariate results will, of course, be restricted to the behavior of the individual variable involved, since we are well aware that one cannot extrapolate joint effects from a combination of significant univariate results when the overall multivariate F is negative.

e) Quality/Quantity issues and the grouping of variables

Previous analyses have given us grounds for looking not only at different kinds of contrasts, but also at different assortments of qualitative and quantitative variables in successive MANOVAS. Thus, some of the analyses below will group the untransformed measures of basic Quality from STL and LL, while others will combine the ratio measures of Efficiency in separate analyses, and still others will omit both measures of raw quality and quantity but include the regression measures of predicted Quality. The main aim of these alternative analysis strategies is to assess the importance of the different roles these measures play in clarifying the direction and magnitude of training influences.

We will also take advantage of the Step-down F test provided in the MANOVAS to order variables for empirical, substantive or intuitive reasons in order to ask questions about variable priorities, e.g., what is the status of Quantity once the effects of Quality are removed?

iii) Results

a) A General Test of the Model

A preliminary analysis was made of the four variables (one from each domain) which were uncovered by the pre-test Discriminant Function analyses, which showed them to be very discriminating representatives of their respective domains. Two of these measures are not summary measures, but are sub-task scores, and this is one of the reasons why this group of variables did not end up in the final set of 15 summary variables chosen for the confirmatory analyses. However, these four merit a brief inclusion here because of their status and because their MANOVA provides a very general test of the model as a whole.

The Perceptual measure is a summary measure of total number of choices across tasks. The Cognition variable is the '3-score' measure of the number of non-egocentric responses. (This is the sub-score that correlates very highly with COGSUM, the summary measure for the Cognitive domain which appears in the main MANOVAS.) The STL variable is the composite measure of Competence (Table 29) and the LL variable is the sub-task score on the Discriminability Index for the Picture Boxes task. This latter task was the hardest in the LL battery and for this particular type of score the post-training results were in the predicted direction, showing E > C and Group III > II > I. (Table M1)

SIMPLE CONTRASTS:

The set of Simple Contrasts was applied to these four measures in a MANOVA. Contrasts D and E were not significant, but Contrast F (group III vs. IV) produced a significant joint effect ($F = 3.08$ (4/42), $p < .03$) (Table M1). All the variables except Perception showed significant univariate effects here also but, as might be expected from the exploratory data, the direction of the effect is different for the STL variable (III(5.91) < IV (15.24), $p < .089$) compared to Cognition (III(1.73) > IV(0.46), $p < .079$) and LL (III(5.27) > IV(2.54), $p < .02$).

This result thus provides some support for the contention that communication training has a general inter-domain effect at this level of comparison of Lo/Hi trainees versus Control Ss. The effects are mediated mainly through the LL and Cognition domains, as might have been predicted from the previous analyses of other summary variables. With regard to the relative impact of training on the different domains, the Step-down F tests the variables from top to bottom in increasing order of importance as indicated by the Discriminant Function analysis, with the strongest variable (LL Discrimination) entered last. These F-test results indicate that the Cognitive and LL measures showed the strongest effects of training, creating greater gains in performance among the Low Ss paired with High Competent peers.

HELMERT CONTRASTS:

An additional set of Helmert Contrasts were run on these same four variables to check for further multivariate effects for training and pairing. Contrasts A and C produced no significant joint effects, but the orthogonal Contrast B (group I (Lo/Lo) vs. groups II (Lo/Mod) + III (Lo/Hi)) produced a significant Multivariate effect ($F = 2.74$, $p < .04$) for the influence of Pairing on test performance among these experimental groups (Table M2). The univariate F tests show that the major contributors to this effect are the Cognitive ($p < .01$) and LL ($p < .05$) measures. These clearly showed that the two groups (II and III) with the more competent peers made greater gains on these variables after training than the Lo/Lo Ss who experienced "zero modeling".

We spoke earlier about training effects being "mediated" from one domain of competence through another. We can now point to the above result as an indication of ability being mediated from one student to another. This finding has implications for our notions of how communicative competence is "learned".

With regard to the role of the STL domain in this general test of the model, we think that the influence of Training on STL is not appropriately evaluated by the kind of contrast (III vs. IV) that was discussed above. This is because of the atypical performance of group III in the post-training assessments, as reiterated before. What is needed are more appropriate tests within the STL domain which do not mask the real gains made by groups I and II. Alternative analysis strategies for dealing with this problem will be discussed below.

b) Perception and Cognition

Turning now to the major summary variables in the analysis, we will first examine differences in unadjusted post-test scores in the Perceptual and Cognitive domains. The correlational analyses showed there to be a significant increase in relations between these two over time, so a case can be made for including both domains in the same MANOVA.

HELMERT CONTRASTS (1):

The set of Helmert Contrasts were applied to the two Perceptual measures and the Cognitive summary measure and the results for the three Contrasts are shown in Table M3. There were no significant effects of Training on these measures, and among the tests of Pairing effects only Contrast 'B' gave a significant multivariate result

($F = 2.34$, $p < .08$) showing groups II + III > I). While this indicates that there is a significant joint effect between the two domains on these measures, the Unova results would suggest that the strength of this Pairing effect is largely contributed by COGSUM ($p < .01$). As will be remembered from the data in Table 14, the scores on this Cognitive measure were perfectly ordered III > II > I, and the present result confirms the significance of this Pairing effect and of the greater reduction of egocentricity among Low SS paired with more competent partners.

What about Training effects? The lack of a significant result on these measures tends to confirm the previous findings that the Perceptual measures produced no clear pattern of results in this respect. Yet one might still ask, as was done in the correlational analyses, if there might not be a joint effect of training linking the Perceptual domain with a perceptually "loaded" variable in the language domain, like the Discriminability Index. And since the latter measure seemed to have not a direct but an indirect relation with Perceptual Accuracy mediated via Cognition, it was decided to repeat the above analyses on Perception and Cognition together, with the inclusion this time of the LL Discrimination measure.

HELMERT CONTRASTS (2):

The results of the Helmert tests on the post-test scores for these measures are shown in Table M4. This time there is a significant joint effect for Training (Contrast "A", $F = 3.06$, $p < .03$) on the three domains, and the Trainee Ss differ very significantly from the Control Ss on the LL Discriminability measure (Univariate $F = 11.58$, $p < .001$). This latter result is the major reason for the significant multivariate effect, but this joint effect still at least implies some inter-domain influence, however indirect it may be. [The Step-down F tests for Contrast A in Table M4 show that the LL variable effect is still as strong when the effect of Perceptual Accuracy is removed. This would confirm what was noted in the earlier correlation analyses, namely, that there is no direct relation between these aspects of Perceptual and LL functions, so the hypothesis of an intermediary relation still stands.]

Contrasts B and C in Table M4 produced no significant results for Pairing effects when the LL variable is included in the analysis. The only significant univariate effect is for Cognition in Contrast B, but this is the same result that was discussed above in the first MANOVA run for Perception/Cognition.

c) Cognitive Gains

The exploratory analyses showed significant differences for Training and Pairing in absolute levels of post-training performance on the Cognitive summary measure. Those analyses also suggested that there were equally important patterns of score changes over time for this variable that ought to be evaluated, since they also seemed to conform to the original predictions for Training and Pairing effects. Accordingly, a supplementary analysis of post-training gain scores in the Cognitive domain was made in conjunction with the main confirmatory MANOVAS.

As part of a set of Helmert tests of some sub-task measures, orthogonal comparisons of group means for COGSUM gain scores were produced, and the Univariate F tests for both Training and Pairing effects were significant. For Contrast A, the Experimental groups made greater gains after training than the Control group ($F(1/45) = 3.04$, $p < .088$).

Contrasts B and C produced a nicely balanced set of Pairing effects. In Contrast B, both of the high-scoring groups II and III improved significantly more than the Lo/Lo Ss in group I ($F = 4.88$, $p < .03$). Groups II and III, on the other hand, were not significantly different from each other. Further Simple Contrasts of individual training groups with the control group showed both group II ($II > IV$, $F = 5.33$, $p < .02$) and group III ($III > IV$, $F = 3.36$, $p < .07$) to have made significantly greater gains in cognitive functioning than the Control group.

When the results of these gain score analyses are combined with the earlier significant results demonstrated with the unadjusted post-test scores, they show that the influence of communication training on this cognitive domain is both clear and strong, in both absolute and relative terms. The interaction of this significant domain with the language domains will be studied in further analyses below.

d) The Spatio-Temporal Language Domain (STL)

As was seen in the exploratory analyses of the STL post-test scores, the unpredicted, atypical performance by the Lo/Hi group of Ss had greatly attenuated the mean score for the three experimental groups taken together and thus vitiated the planned comparisons designed to test training and pairing effects. Because of the asymmetric score pattern produced in the post-test, the planned contrasts effectively masked what appear to be real training effects for groups I and II.

HELMERT CONTRASTS:

Even though inappropriate, the set of Helmert Contrasts were applied to the post-training data for a group of four STL Summary variables for the sake of demonstrating the obvious. The four variables were — Quantity, the log transform measure of Quality, Organisation and the Accuracy of Prepositional Usage score. As expected, none of the three Contrasts produced any significant multivariate effects.

ARBITRARY CONTRASTS:

It was therefore decided to construct a more realistic set of contrasts, which would make more meaningful comparisons of the apparent differences between the Control group and groups I and II. It was further decided that these contrasts would be applied not to the same post-test scores as previously, but to the pre-post gain scores for the STL variables, since these measures of performance change seemed to provide even stronger evidence of experimental vs. control differences. Accordingly, a set of three Arbitrary Contrasts was constructed: I + II vs. IV, III vs. IV; and I vs. II. Only the first of these was of interest, and it could reasonably be claimed that this was the independent comparison in the set. The results from the second and third contrasts were ignored. The first Contrast compared the mean gains of Training groups I and II with those of the Control group. It operated on the gain scores of a group of five STL major variables, and both the mean scores and the results of the multivariate analysis can be seen in Table M5. The five measures were: 1) basic propositional Quality, 2) message Inefficiency, 3) the simple ratio measure of message Efficiency, 4) message Organisation and 5) overall STL Competence.

RESULTS OF GAIN-SCORE ANALYSES:

As predicted, Table M5 shows significant multivariate and univariate effects of Training in favor of the gains made by the two treatment groups after participation in the program. The mean scores in Table M5 show that, to a greater extent than the Control Ss, the Lo/Lo and Lo/Mod groups have improved the overall quality of their propositions, have reduced the inefficiency and increased the efficiency of their messages, have improved the organisational structure of their messages, and generally become more competent overall in this type of STL Usage. The joint Manova effect reflects this overall superiority of the experimental Ss, while the Univariate F tests show that the changes for three of the five variables are significant — Efficiency ($p < .04$), Organisation ($p < .06$) and Competence ($p < .01$).

Out of curiosity, another analysis was run applying the planned Helmert Contrasts to the same STL gain scores. None of the three comparisons produced significant multivariate results (as could only be expected), but Contrast C II vs. III) produced significant univariate results again for STL Organisation ($II > III$, $F = 3.8$, $p < .058$) and Competence ($II > III$, $F = 5.92$, $p < .02$), simply highlighting once more the discrepancy between the performance of the Lo/Hi trainees and that of the Lo/Mods.

DISCUSSION:

There is no need to comment on Pairing effects here since the unpredicted pattern of results within the Training groups (i.e., $I > II > III$) destroyed the rationale of the original within-group comparisons. We can only direct the reader back to the alternative explanations of this phenomenon, which were offered earlier in this report. What is worthy of comment, however, is the significant training effect obtained with an alternative analysis strategy, and what this implies for the kinds of scoring variables that were involved. All three of the STL variables that were significant in the Univariate analyses in Table M5 are all composite measures of one sort or another, involving formulae that combine quality and quantity, or sequence and organisation of propositions or, in the case of the composite measure of competence, combining quality, quantity and organisation in one measure. This kind of result again speaks to the importance (indeed the necessity) of language measures whose scoring systems capture both the Form and the Content of message construction. In particular, the composite measure of STL 'Competence', which proved to be such a differentiator of training effects in the above analysis was also the variable identified for its power in the pre-test Discriminant Function analyses. It would thus appear to be a very representative indicator of both ability and learning for this type of language function.

e) The Logical Language Domain (LL)

The exploratory analyses of the major LL variables did not show the kind of aberrant performance among the groups that was encountered in the STL domain, and predicted patterns and trends in the scoring seemed to be clearly enough established through the use of the unadjusted post-test scores without recourse to alternative analysis strategies. Accordingly, the Helmert contrasts were still the multivariate tests of choice for these LL data, and the prediction was that they would produce a straightforward set of significant results for training effects.

HELMERT CONTRASTS:

This prediction was confirmed by the results of the MANOVA which can be seen in Table M6. Four summary LL variables were entered in the analysis -- Organisation, Efficiency, Quantity, and Discrimination (Tables 32 - 37). The first Helmert Contrast 'A' produced a highly significant joint effect ($F = 3.34$, $p < .018$) showing the three training groups together to be scoring higher after training than the Control Ss all across the LL domain. With the exception of Efficiency, all of the LL variables produced very significant Univariate results also. However, the Step-down F tests are relevant here also, for the four variables were entered into the analysis in an order based on their empirical relations shown in previous analyses. This allows us to see some fascinating results from the Step-down F tests: Efficiency per se with the effects of Organisation taken out, starts to become significant ($p < .08$), while Quantity, with the effect of Efficiency removed drops out!

The remaining two tests of Pairing effects within the experimental groups (Contrasts B and C) produced no further significant results. This corroborates the conclusion from the exploratory analyses that there are no clear-cut effects of peer-modeling in this domain. Nevertheless, we can point out that on significant and powerful measures like LL Organisation and the Discriminability Index, the groups with the more competent peer models (II and III) score higher than the Lo/Los (group I).

DISCUSSION:

In conclusion, this first confirmatory analysis of the independent LL domain shows it to produce a unique set of results attesting to the effectiveness of training in improving message fluency, organisation, discrimination quality of propositional content, and to a lesser degree, overall message efficiency. The MANOVA results also provide some indications that the different measures of categorical Quality, Efficiency and Organisation are all useful differentiators of message construction ability in this domain, and allow us to look at complementary aspects of this process. Further analyses below will examine Quality/Quantity relations in LL and STL, as well as the relations between Cognition and the two langauge domains.

f) Cognition and Language

The correlational analyses suggested that the summary measure of cognitive perspective-taking ability had different relations with each of the language domains, and that this measure of egocentricity and decentering ability had stronger and more direct relations with the Logical Language measures. The original model underlying the study had assumed positive relations between cognition and language at a general level and open-ended expectations were entertained regarding changes in these relations after training. In view of the more differentiated picture of the cognition-language relations provided by the exploratory analyses, a pair of confirmatory MANOVAS were run which included the Cognitive summary measure in the groups of STL and LL variables studied in the earlier runs.

HELMERT CONTRASTS (1):

In the first MANOVA, Cognition x STL relations were examined. The summary measure, COGSUM, was entered first in the group of variables, which included in order, STL Organisation, the transformed measure of Efficiency, and the vocabulary measure of Prepositional Usage (see Tables 14, 17, 22, 26). The results of the three Helmert Contrasts of these data are shown in Table M7. There are no significant effects for Training (Contrast 'A') and of the two tests for Pairing effects only Contrast 'B' is significant. This shows that groups II and III are significantly different from group I (F(4,42) = 2.81, $p < .037$) among the trainee Ss. The meaning of this joint effect for post-test scores should probably be interpreted with caution, for in the case of the two significant univariates here (COGSUM and STL Organisation) the pairing effects are in entirely opposite directions. For the Cognitive measure ($III > II > I$, $p < .01$) the direction of effect is as predicted and as indicated in earlier analyses. For the message Organisation variable, on the other hand, the direction ($I < II < III$, $p < .077$) confirms the general pattern of upside-down results we have already encountered for the STL domain as a whole. Taken with the significant gain-score results for STL Training effects reported above, the present result simply confirms that there is a differential effect of pairing on the raw post-test scores of these three Trainee groups and leaves the causes open to interpretation.

HELMERT CONTRASTS (2):

The results of the second confirmatory Manova on Cognition x LL are shown in Table M8. The same four LL variables as above were entered into this analysis along with COGSUM, i.e., LL Organisation, Efficiency, Quantity and the Discriminability Index (see Tables 32 - 37). This time the directness of the relation between these two domains is confirmed. Helmert Contrast 'A' shows a significant joint effect for Training (Experimental > Control, F = 2.61, $p < .039$). There is no significant univariate result for cognition in this Contrast, but three of the LL variables do show such effects, as they also did in Table M6 above. Of the remaining two Contrasts, 'B' shows a significant multivariate effect for Pairing ($II + III > I$, F = 2.08, $p < .087$), but the bulk of this appears to be carried by COGSUM, which has the only significant univariate F ($p < .01$) for this contrast.

The results for this MANOVA repeat the pattern demonstrated in Table M6 above for the LL domain but, in addition, they also confirm the involvement of Cognition in the Training and Pairing effects for this domain. The notion of a direct relationship between these two domains is further reinforced by the results of the Step-down F test for Contrast 'A' in Table M8, which shows that when the effects of Cognition are removed, there is some weakening of the univariate effects for the significant LL variables.

DISCUSSION:

Communication training, therefore, has positive effects for Low competent Ss on at least two of our domains of competence, benefitting their cognitive functioning and LL usage at one and the same time. The effects of this kind of oral language program thus appear to generalise to areas of cognitive functioning and the significant Pairing results suggest that the vehicle of transmission appears to be exposure of the Low competent Ss to more able peer models. We are left with the intriguing problem of deciding whether the Moderate and High competent models are more effective because they have more cognitive ability, or more LL ability, or both.

The relations between Cognition and the STL domain are less strong, direct, or definite, when measured by this type of post-test score analysis and these kinds of contrasts. The one significant joint effect for Contrast 'B' suggests a putative relation at the level of Pairing, but the significance of this is dubious, especially in the light of the previous correlational analyses showing the lack of relation between COGSUM and STL in general, and between COGSUM and STL Organisation in particular. We are left, therefore, to fall back on the gain-score analyses to understand the relation between training and STL performance. Future research should probably look at MANOVAS that group STL variables with Perception measures, and which provide gain-score analyses of Cognitive with STL measures, with more appropriate contrasts.

g) The two Language Domains and the Quality/Quantity Issue

Having looked at STL and LL domains separately we can now assess the interaction or interdependence between these two domains which were the major targets of the training program. In the context of these joint analyses we will also evaluate the different roles played by the different measures of propositional Quality and message length or Quantity, in an attempt to determine which provides the clearest picture of training effects and inter-domain relations.

Three types of MANOVAS will be done. In the first, STL and LL variables will be grouped together using the transformed ratio measures of Efficiency which are adjusted for variations in message length in both domains. In the second, the same language measures will be used, but this time the Efficiency measures will be replaced by the raw, untransformed, summary measures of propositional Quality uncorrected for length, from both domains. In the third MANOVA, the measure of propositional quality for both domains will be the regression transformation measures of Predicted Quality.

The groups' means for all of these variables can be seen in Tables 17-30 and 32-38.

1) STL x LL x Efficiency Scores

This MANOVA applied the Helmert Contrasts to the unadjusted post-test scores for the variables in both domains, for the sake of comparison and consistency. There were four LL measures as before (Organisation, Efficiency, Quantity and Discrimination) and three STL measures (Prepositional Usage, Efficiency and Organisation). The results are shown in Table M9. There is the glimmer of a significant joint effect for Training ($F(7,39) = 1.88$, $p < .099$) but the reasons for this seem to be due to the significant univariate effects for the same LL variables that were significant before. Neither of the other two Contrasts produced any significant results.

It should be remembered, of course, that we have previously argued that the helmert Contrasts are inappropriate for this type of score for the STL measures. With this caveat, it would appear that at this level of analysis, at least, and with these particular assessments of message efficiency, the two domains retain their relative independence.

2) STL x LL x Raw Quality/Quantity Scores

The same MANOVA was repeated, this time with two different STL measures (Quantity and the basic summary measure of Quality) along with Organisation and Prepositional Accuracy, and with the LL summary measure of propositional Quality accompanying the other three LL measures.

The results of this MANOVA are in Table M10. This time there is no significant multivariate result for Training ($p < .19$), but there are some interesting univariate results in this Contrast that are worth commenting on. In addition to the usual significant F for LL Discrimination, Organisation and Quantity, there is an important result for the raw Quality measure in LL that confirms predictions based on the previous exploratory analyses. This shows that all three Training groups scored significantly higher ($p < .04$) than the Control Ss in amount of positive propositions produced after training. As regards the relation of this measure to message length, it is interesting to note in the Step-down F that when the effect of LL Quantity is removed the significance of this Quality measure drops out. Also noteworthy is the effect of the STL measures on LL Quantity in the Step-down F test, so perhaps we do have a relation between the two domains at the level of Fluency, as was hinted by the prior correlational analyses.

There were no significant effects of Pairing in the other two Contrasts, but we can draw attention to the significant univariate effect for STL Organisation ($p < .077$) in Contrast B in Table M10 for a special reason. The Step-down F relation between this measure and STL Quality here confirms the integral relationship between these two aspects of message formulation that was demonstrated in the earlier correlational analyses.

3) STL x LL x Predicted Quality

The final MANOVA in this set used the predictor measures of Quality with the effect of quantity removed by regression for both domains, along with the usual measures of Organisation, Discrimination, etc. In all, there were three STL and three LL measures. Table M11 shows the results.

This time there is an indication of a joint inter-domain effect for the test of Training (Contrast 'A', $F(6,40) = 2.01$, $p < .087$), although this seems to be largely due to the action of the two LL variables (Organisation and Discrimination) which have the only significant univariate effects for this contrast. The remaining two contrasts produced no significant results for Pairing effects.

DISCUSSION

Among the three MANOVAS, therefore, the analyses which produce any signs at all of a small, joint effect are the two in which some attempt is made to control for the Quality/Quantity relationship, either through the use of the log transformed or the regression transformed measure of Efficiency. What these results seem to be confirming is that competence is not just fluency in generating propositions. There are factors of quality and efficiency of message construction involved, and this is a point that was made earlier when we pointed to the importance of the STL composite measures of efficiency and competence which took into account qualitative and organisational aspects of message formulation as well as quantitative ones. The same point could be emphasised also in the case of the significant results in the LL domain in Table M6 where the Step-down F results indicated that the more important variables are those which measure not just message length, but message organisation, efficiency and discriminability.

These MANOVA results ought to be interpreted in the context of the prior correlational and factor analyses which showed, paradoxically, a strong positive relationship between raw quality and quantity of propositions, but an increasingly negative relation between Efficiency and length over time. To resolve this paradox, a general interpretation of the MANOVA results can be offered here. What the inverted Efficiency/Quantity relation may be telling us is that, after Training, Efficiency increases at the cost of length and that fluency per se is no longer crucial. Yes, propositional quality goes up and yes, length goes up, because the child gets a chance to say more positive things with longer utterances. From the point of view of learned competence, then, efficiency makes better use of message length. This seems to be particularly true for the significant effects of training in the LL domain.

h) GRAND MANOVA

A final MANOVA was run lumping together all the summary variables from the four domains in one test of inter-domain effects. Eleven variables were entered into the multivariate analysis, grouped and ordered for substantive reasons. The first two variables are the Organisational measures from both language domains, the fourth and fifth are the Quantity measures for LL and STL, the sixth and seventh are the language Efficiency measures, the 8th, 9th and 10th are the LL Discriminability Index grouped with the two Perceptual measures, and COGSUM brings up the rear.

The results of this Grand MANOVA can be seen in Table M12, using the same set of Helmert contrasts operating on the same kinds of unadjusted post-test scores as before. There were no significant joint effects for any of the three Contrasts, while the only significant UNOVA effects were the ones already seen in previous analyses for Training (LL Organisation and Discriminability) and Pairing (STL Organisation and Cognition). The robustness of the Pairing effect (III > II > I) in the Cognitive domain can be appreciated from its continued strong showing in 11th place in the Step-down F analysis in Contrast 'B'.

On the face of it, these results continue to show within-domain relations to be stronger than those between domains in our model. A more discriminating test of between-domain relations was probably provided by the Discriminant Function analyses in the pre-test Evaluation phase of the study, and the gain-score MANOVA (Table M1) of the most discriminating variables from those analyses showed there to be significant interrelations between at least two domains -- Cognition and Logical Language -- in their responsiveness to Training, at the highest level of peer-modeling. The significant role of these two domains runs consistently through these confirmatory analyses, which generally support the original predictions for these two areas of functioning, for these kinds of experimental comparisons of absolute levels of post-training performance. Running parallel to these findings, but on a separate track, are the set of significant results for the other language domain, STL, where the effect of training is more clearly seen in the amount of change that takes place in competence over time.

It should be noted, of course, that although the post-training analysis strategies in the two domains may differ, they do share a common reference point in the initial equivalence of the experimental groups that was established in the pre-test before the training began. So, whether one speaks of higher or lower levels of post-training performance, or of greater or lesser amounts of gain over time, the baseline zero point for both kinds of comparison remains the same.

1) Summary Discussion of MANOVA Results

Any generalisations about the results of these confirmatory analyses have to be qualified, of course, by the peculiar nature of the STL data which required the use of alternative, unplanned analysis strategies to tease out the effects of training and pairing on that particular domain. Those constraints and caveats aside, the multivariate analyses confirmed many of the patterns of effects espied in the earlier exploratory analyses.

There is, for example, the confirmation of the relative independence of the two language domains. We would stress the word 'relative' here because there are common elements of vocabulary and fluency which were previously seen to link both domains. In areas like these the correlational analyses showed there were positive but low correlations (with r's between .3 and .4) between the two domains, and this level of relationship seems to be confirmed by the suggestions, in the MANOVA results, of a small, joint inter-domain effect of Training on language learning. This effect was most apparent in analyses of message efficiency where a covariate correction was made in both domains for variations in message length.

The strongest language relations are, of course, still within-domain, and the multivariate analyses showed that each domain responded in its own unique way to the influence of training and peer-modeling. In STL the exploratory univariate analyses showed that the post-test scoring trends favored 2/3 of the trainees over the controls at least half of the time, but that there were even more important trends, on virtually all of the measures, for the relative gains made over time by these same groups of Low competent Ss. It is not surprising, therefore, that the gain-score MANOVAS proved to be more appropriate for uncovering important Training and Pairing effects among this group of STL variables. These showed that composite measures of message construction ability, which captured important structural and substantive features of utterances in a single score, were the most effective indices of learning among the less competent Ss in two out of the three Training groups.

Within the LL domain, the exploratory analyses showed that there was no need to change the planned comparisons of treatment effects because, on all of the major summary measures, the Experimental Ss outscored the Controls in absolute level of post-training performance. This superiority was confirmed by the MANOVA, which also showed that Pairing effects were less direct in this domain. These appeared to be mediated by the Cognitive domain which seems generally to have a strong, positive relation with logical language functioning.

The relation between Cognition and STL still continues to be less clear or direct. The hypothesis of an indirect, mediated relation between Cognition, Perception, and STL can still be offered, but the present MANOVAS, with their particular kinds of experimental contrasts, did not provide a proper test of this assumption which will need to be studied in further research.

One indirect relationship which was tested is that between Perception and LL, and the significant joint effect uniting Perception, Cognition and LL (Table M4) continues to give credence to the notion of an indirect, mediated relationship here. In other analyses, this particular set of summary Perceptual measures continued to produce the same uninspiring results they did in the exploratory analyses. There is still a need in this domain for a good composite measure which effectively combines Speed and Accuracy of choice into a single discriminating variable.

Although Training effects were always strong in the LL domain, Pairing effects were not always as clear-cut on all of the variables. There are still indications, however, in the confirmatory analyses, that both Training and Pairing effects do show up on more powerful and more critical LL variables, such as the Discriminability Index. It will be remembered from the MANOVA on the Discriminant Function variables that one of the strongest measures in the competence battery was the Discrimination score from the hardest LL sub-task, Picture Boxes. On this measure the predicted trends were obtained showing Experimentals > Controls, and groups II and III higher than I. The same pattern of results was obtained for one of the most significant measures in these confirmatory analyses, the summary Discriminability score from all of the LL sub-tasks. One interpretation of these patterns of results is that communication training produces real learning in the LL domain by improving Low subjects' performance significantly on the harder tasks, and achieves this partly as a function of the contribution made to this learning by the more competent peer-models.

Finally, the one result that stands out in all of these analyses is the significant change that has taken place in the cognitive perspective-taking ability of the Low competent Ss, both in absolute level of performance and in the amount of development over time. The pedagogical value of the training program would be vindicated by the Cognitive results alone.

III. 4. SUMMARY AND CONCLUSIONS

This research study has been concerned with three perennial topics in applied psycholinguistic research -- Evaluation, Intervention and Learning. It has made significant contributions to each of these areas by developing original measurement techniques, by creating new strategies and methods for fostering language learning, and by systematically and empirically assessing the various effects of intervention on the development of communicative competence in less able pupils.

The theoretical model which underlies the whole study has also undergone modification and revision in the process of its practical application. It has become more differentiated, with four major components now instead of three. The Language domain has split into two important functional areas, Spatio-Temporal Language usage and Logical Language usage, and this has important implications for the representativeness of assessments of language abilities of children in the future. Previous studies in this area have grossly underrepresented the child's pragmatic capacities by concentrating on limited aspects of only one area of language functions at a time. The results of this study now call for assessments of communicative competence that are more differentiated and relative, and less one-sided and myopic.

The theoretical model also requires modification to take account of the dynamics that operate between its different domains or processes, since Perception and Cognition interact in different ways with Language while, within the latter domain, the two major functions appear to operate in an atmosphere of relative independence. Important relations have been identified between Cognition and Logical Language, which amount to a demonstration of the effects of a language learning program generalising to areas of non-language function. Furthermore, the important role of peer-models in mediating such effects has important implications for the nature of language learning environments in the classroom.

The educational implications of this study have already been discussed in detail in the concluding sections of Part II of this report, so the reader is referred to that discussion. We prefer to deal here with the important conceptual and methodological contributions of the study, especially in regard to the Measurement and Evaluation of communicative competence. The theoretical model has successfully bridged the gap between laboratory and classroom in this area by helping us to conceptualise practical ways of describing communication behavior and by providing a framework for developing the necessary instruments for analysing, measuring and evaluating that behavior. We feel that the Propositional Analysis scheme developed in this study, along with its taxonomy of oral message characteristics, can make an original contribution to the problems of describing, scoring, and categorising children's pragmatic utterances in ways that are meaningful for developmental comparisons, as well as for educational diagnosis and prescriptive planning. The scoring concepts involved in this analysis scheme for LL and STL functions in the older child also find an echo, for example, in some of the current psycholinguistic research on the acquisition of stative and active verb forms in infancy and early childhood (Bloom, Lightbown and Hood, 1975; Bloom, Hood and Lightbown, 1974).

In addition, the battery of perceptual, cognitive, and verbal tests developed for the Evaluation Phase of the study show a lot of psychometric stability as measurement instruments. As the various analyses showed (Discriminant Functions, Canonical Correlations, Reliability, etc.), the items and tasks within each of the language domains appear to hold together fairly well, while the test-retest data indicate a good deal of consistency and reliability of measurement. The Cognitive Scale also shows good external validity in its high correlation with the already established Piagetian Mountain Perspective Task. The competence battery as a whole can reliably and meaningfully describe and differentiate levels of pragmatic-linguistic ability in schoolchildren. The utility of the battery could be improved even further by using the empirical data of this study to boil down the tests to a smaller, optimal set of tasks and measures which could be conveniently used by teachers, clinicians and researchers. In addition, the Language tasks in the battery are also interesting enough and varied enough in themselves that they could also be used as teaching or training materials if the practitioner so desired.

In conclusion, this study has created a number of original products for both testing and training communicative competence in schoolchildren, and it has produced a variety of original findings regarding the comparative diagnosis of pragmatic-linguistic abilities and the effectiveness of novel pedagogical treatments on these competencies.

III.5 Recommendations

1. This study focused on the assessment and training of expressive language abilities of less competent children due to the critical importance of such skills and to the fact that a well-differentiated scheme of analysis of pragmatic utterance forms in the school-age child is not available in the research literature on referential and other forms of communication. But the kinds of questions and issues raised in the course of this project can also be applied to the receptive side of the process and companion research needs to be initiated concerning schoolchildren's listening or decoding abilities. There is a current national concern in education with minimal competency in speaking and listening, but not only is there a lack of empirical data on which to base programs to evaluate and assess listening competencies, there are even fewer good theory-based or data-based programs available for training listening competency in school children. A coordinated program of basic research on speaking and listening is badly needed.

For the developmental psycholinguist, this study also raises researchable questions about the development of pragmatic listening competency and its coordination with speaking competencies in the growing child. We have already mentioned, for instance, meta-message or meta-linguistic aspects of utterance interpretation and comprehension. There are other issues, too -- e.g., is a good encoder also a good decoder? How can a listener understand and correctly respond to inadequate messages? Conversely, why is an adequate message incorrectly decoded? Similar research issues have recently been identified by Sinclair (1975) from a Piagetian perspective (cf. also Shatz, 1978).

2. Following the example of the Bullock Report (1975), the integrationist approach favored in this study inevitably raises questions about the relationship between communicative competence, on one hand, and reading and writing on the other. We have already discussed possible relations of oral language learning to the development and teaching of writing ability. This study's demonstration of the relations between Cognition and Logical Language usage have direct implications for future systematic research on the relations between oral language development and literacy. Recent research by Kroll (1978) is a step in the right direction here. In addition, the educational philosophy and technique advocated in Part II of the report could be profitably applied in a similarly structured, intensive writing program.

3. The limited resources of the present study were concentrated on an analysis of fairly large utterance units at the message level. Work remains to be done on more refined, qualitative linguistic features (syntactic, semantic, etc.) of the children's oral communications so that these can be correlated with the other indices of competence developed in this study. A more detailed linguistic analysis of the data already collected would show in what respects the vocabulary and grammatical structure, for example, of the Low competent children differ from those used by the High competent Ss.

4. Further research also needs to be done to clarify the relation between the Perceptual, Cognitive and STL components in the model of communicative competence. The model's effectiveness would be enhanced by exploring in more detail the role of the perceptual component and whether its relation with STL is at the level of visual analysis or visual recall of target features, as well as the nature and purpose of cognitive input to this process.

Methodologically, different analyses strategies should be employed to evaluate these inter-domain relations. A scoring strategy that could be profitably employed in both univariate and multivariate analyses of Training effects would be the use of pre-post gain scores covaried on pre-test starting-level scores. This would be a way of dealing with floor and ceiling effects of hard and easy items on measurement, and it would also eliminate the need to statistically establish initial equivalence among groups.

5. This study pointed up the value of composite measures which are sensitive to relations between structure and content, and between the parts and the whole of a message. We believe that Sinclair's notions of the different types and levels of "knowledge" displayed in a child's utterances provide a useful conceptual framework for further research on the relations between the segments and the meanings of a message. This approach should be used in further basic research on the development of the message construction process in the child and how the ability to use and understand oral messages evolves. One research topic on meta-linguistic or meta-message awareness could be, e.g., how much does a child have to tell you before he or she thinks they have said enough about a situation or an event? Our notion of "First Message" analysis was an attempt to deal with this intriguing pragmatic issue of the child's communicative intent.

6. From a practical point of view, more experimental field-work needs to be done in school systems and classrooms on the developmental sequencing, structuring and refining of the communication curriculum. As was pointed out in Part II of this report, the levels of activities within a game category can be ordered according to logic or complexity. However, this curriculum makes no claims that an optimal, hierarchical order yet exists between game categories, with respect to differential difficulty or developmental appropriateness for children of differing ages, needs, or abilities. It is still an open question, for example, as to whether games like Who Am I?, should be introduced before or after categorising games like Bingo or Go Fish. This, in turn, will depend partly on an individual teacher's preference for advancing through the curriculum successively, one game at a time, or simultaneously, pushing game activities on several fronts at once. Coordinated feedback from teachers using the program would provide valuable comparative data that would be helpful for

refining the curriculum and establishing empirical, developmental guidelines for a timetable of application for the program.

Establishing a curriculum sequence was beyond the scope and resources of the present project. The Training Phase of the study was begun with a small core curriculum in place, but most of the game material was developed as the project went along. We never had the luxury of being able to sit back and manipulate the sequencing of activities. However, now that a completed curriculum exists, teachers will be able to experiment with different strategies and timetables for implementing the curriculum over the school year.

At the practical level, therefore, the rare opportunity exists to work with and further refine a communication curriculum whose rationale and effectiveness are already solidly based on theory and empirical research.

7. A final recommendation to future users of this communication curriculum deals with the issues of "balance" and "integration". With regard to establishing balance in the curriculum, it seems in retrospect that our experimental curriculum may have been a little imbalanced in having more activities of the LL type than of the STL variety. It is difficult to assess the impact of this on the Training data, especially since the game types in both domains frequently used the same situations (screened vs. face-to-face, dyadic vs. small group, etc.) and materials (toys, abstract or concrete objects, etc.) while varying the communication purpose. Future users of the program should try to achieve a balanced curriculum with equal emphasis on STL and LL activities. STL usage could be promoted, for example, by including more games along the lines of the Story Ordering game.

The issue of integration deals with the relation of this communication program to the whole of the child's elementary curriculum. We end this report the way we began it by citing Bullock's recommendation for "language across the curriculum". We urge teachers who plan to use this program and to evaluate its outcomes in their classrooms to integrate this use and evaluation with the promotion and evaluation of achievement in other areas, such as cognitive development, reading, math, etc. The developmental approach, the long-term view, the concern with process rather than product, and solid evaluation -- all of these make a worthy recipe, not just for language education, but for the education of the whole child.

BIBLIOGRAPHY[1]

Austin, J.L. How to do things with words. Harvard Univ. Press: Cambridge, Mass. 1970.

Baldwin, T., & Garvey, C. Studies in convergent communication:II. A measure of communication accuracy (Report #91). Center for the Social Organisation of Schools, Johns Hopkins Univ.: Baltimore, Md., Nov. 1970.

Baldwin, T., McFarlane, P. & Garvey, C. Children's communication accuracy related to race and socioeconomic status. Child Development, 42, 1971.

Bates, E. Peer relations and the acquisition of language. In M. Lewis & L. Rosenblum (Eds.), Friendship and Peer Relations. Wiley: N.Y.. 1975.

Bates, E. Language and Context. Academic Press: N.Y. 1976.

Bernstein, B. A sociolinguistic approach to socialisation. In F. Williams (Ed.), Language and Poverty. Markham Publish. Co.: Chicago, 1970.

Bock, R., & Haggard, E. The use of multivariate analysis of variance in behavioral research. In D. Whitla (Ed.), Handbook of Measurement and Assessment in the Behavioral Sciences.:Addison-Wesley, Mass. 1968.

Bock, R. Multivariate Statistical Methods in Behavioral Research. McGraw-Hill:N.Y., 1975.

Brown, R. Social Psychology. Free Press: N.Y., 1965.

Brown, R. A first language. Harvard Univ. Press: Cambridge, Mass. 1973.

Bullock Report: A Language for Life. Report of the British Govt. Commission of Inquiry. Dept. of Education & Science: H.M.S.O., London, 1975.

Cazden, C. The situation: a neglected source of social class differences in language use. J. of Social Issues, 26, 1970.

Cazden, C. Child language and Education. Holt, Rinehart & Winston: N.Y., 1972.

Cazden, C., John, V., Hymes, D. (Eds.) Functions of language in the classroom. Teachers College Press: N.Y., 1972.

Chomsky, C. The acquisition of syntax in children from five to ten. M.I.T. Press: Cambridge, Mass., 1969.

Chomsky, C. Linguistic developemnt in children from six to ten. Final Report, U.S. Office of Education Project #9-A-055, Harvard Grad. School of Education, 1971.

---

1. See Addenda at end of references.

Chomsky, C. Stages in language development and reading exposure. Harvard Educ. Review, 42, 1-34, 1972.

Chomsky, N. Language and Mind. Harcourt, Brace & World: N.Y., 1968.

Cole, M., Gay, J., & Glick, J. The development of communication skills among Liberian tribal people. Paper presented at the Society for Research in Child Development Meetings, Santa Monica, California, March 1969.

Cooper, C.R., Lopez, S.A., & Marquis, A. Children's discourse during peer learning in Experimental and Naturalistic situations. Discourse Processes (in press).

Coulthard, M. An introduction to Discourse Analysis. Longman: London, 1977.

Cowan, P.A. The link between cognitive structure and social structure in two-child interaction. Paper presented at the Society for Research in Child Development Meetings, March, 1967.

Chandler, M., Greenspan, S., & Barenboim, C. Assessment and training of role-taking and referential communication skills in institutionalised emotionally disturbed children. Developmental Psychology, 10, 546-553, 1974.

Dickson, W.P., & Patterson, J. Criteria for evaluating curriculum materials which use referential communication activities to teach speaking and listening skills. (Working Paper #273). Wisconsin Research and Development Center for Individualised Schooling: U. of Wisconsin, Madison, 1979.

Dickson, W.P., & Moskoff, M. A meta-analysis of Referential Communication studies: a computer readable literature review. (Theoretical Paper #83). Wisconsin R & D Center for Individualised Schooling: U. of Wisconsin, Madison, 1980/

Dore, J. Holophrases, speech acts and language universals. J. Child Language, 2, 21-40, 1975.

Dore, J. Transitional phenomena in early language acquisition. J. Child Lang., 3, 12-28, 1976.

Fillenbaum, S. Psycholinguistics. Annual Review of Psychology, 22, 251-308, 1971.

Finn, J.D. A General Model for Multivariate Analysis. Holt, Rinehart & Winston:N.Y. 1974.

Flavell, J.H. Role-taking and communication skills in children. Young Children, 21, 3, 1966.

Flavell, J.H., et al. The development of role-taking and communication skills in children. Wiley: N.Y., 1968.

Furth, H., & Wachs, H. Thinking goes to school. Oxford U. Press: N.Y., 1975.

Garvey, C. Investigation of child-adult differences in problem-solving communication. Final Report, Center for the Social Organisation of Schools, Johns Hopkins University, Baltimore, Md., Feb., 1972.

Gleason, J.B., & Weintraub, S. Input language and the acquisition of Communicative Competence. In K. Nelson (Ed.), Children's Language, Vol.I. Gardner: N.Y., 1978.

Glucksberg, S., & Danks, J. Experimental Psycholinguistics. Erlbaum: N.J., 1975.

Glucksberg, S., Krauss, R., & Weisberg, R. Referential communication in nursery school children: method and some preliminary findings. J. Exptl. Child Psych., 3, 333-342, 1966.

Glucksberg, S., Krauss, R., & Higgins, E.T. The development of referential communication skills. In F.D. Horowitz (Ed.), Review of Child Development Research, Vol. 4. U. of Chicago Press: Chicago, 1975.

Grice, H.P. Logic and Conversation. In P. Cole & J. Morgan (Eds.), Syntax and Semantics, vol.3: Speech Acts. Academic Press: N.Y., 1975.

Grotberg, E.H. Review of Research 1965-1969. Project Headstart, U.S. Office of Economic Opportunity, June, 1969.

Halliday, M.A.K. Explorations in the Functions of Language. Arnold:London, 1973.

Halliday, M.A.K. Learning how to mean. Arnold: London, 1975.

Heider, E.R. Style and effectiveness of children's verbal communications within and between social classes. Doctoral dissertation, Harvard University, 1968.

Higgins, E.T. Communication development as related to Channel, Incentive and Social Class. Genetic Psych. Monographs, 96, 75-141, 1977.

Huck, S., & McLean, R. Using a Repeated Measures ANOVA to analyse the data from a Pre-test-Post-test Design: a potentially confusing task. Psych. Bulletin, 82,4, 511-518, 1975.

Hunt, J. McV., & Kirk, G. Criterion-referenced tests of semantic mastery in school readiness: a paradigm with illustrations. Unpublished report, Univ. of Illinois, May, 1973.

Hymes, D. On communicative competence. In R. Huxley & E. Ingram (Eds.), The Mechanisms of Language Development. CIBA Foundation: London, 1971.

Kagan, J. Reflection-impulsivity and reading ability. Child Development, 36, 1965.

Kagan, J., Pearson, L., & Welch, L. Conceptual impulsivity and inductive reasoning. Child Development, 37, 1966.

Kagan, J., Pearson, L., & Welch, L. Modifiability of an impulsive tempo. J. Educ. Psych., 57, 1966.

Karnes, M.B., & Lee, R.C. *What research says to the teacher in Early Childhood Education*. Monograph of the Council for Exceptional Children: Reston, Va., 1978.

Karnes, M.B., & Lee, R.C. Games as an Instruction Strategy in Early Childhood Education. Paper presented to American Educational Research Association, Annual Meeting, Toronto, Canada, March 28, 1978.

Koen, F. The codability of complex stimuli: three modes of representation. Unpublished paper, U. of Michigan, 1965.

Krauss, R., & Rotter, G. Communication abilities of children as a function of age and status. *Merrill-Palmer Quarterly*, 14, 2, 1968.

Krauss, R., & Weinheimer, S. Changes in reference phrases as a function of frequency of usage in social interaction. *Psychonomic Science*, 1, 1964.

Kroll, B. Cognitive egocentrism and the problem of Audience Awareness in Written Discourse. *Research in the Teaching of English*, 12, 3, Oct. 1978.

Kohlberg, L., & Mayer, R. Development as the aim of education. *Harvard Educ. Review*, 42, 4, 1972.

Lee, W.R. *Language-teaching Games and Contests*. Oxford U. Press: London, 1965.

Leler, H. Language development of socially disadvantaged preschool children. *Project Literacy Reports*, Cornell U., July 1967.

Mandelkorn, B. The development of a communication training technique for young EMRs from age seven to eleven. Doctoral dissertation, Harvard Graduate School of Education, 1973.

Mandelkorn, B., & Corman, L. The effects of two training models on communication-related abilities and intelligence in young IQ-defined EMRs. *Studies in Learning Potential*, 3, 45, 1973.

McCaffrey, A. Communicative Competence and the Disadvantaged Child: studies of the relationship between language models and communication behavior in disadvantaged preschoolers. *Final Report of the Psycholinguistics Project (OEO #B99-4804)*, Harvard Graduate School of Education, September, 1970. (ERIC microfiche no. ED #047 806).

McCaffrey, A. Communicative competence-- how it can be measured and how it can be fostered in young schoolchildren. Paper presented to the *Third International Child Language Symposium*, U. of London, England, Sept. 3-5, 1975.

McCaffrey, A. Parler en classe: une manière d'aborder la communication orale dans les classes élémentaires, de façon non-didactique. *Québec Français*, 25, March, 1977.

McCaffrey, A. From Competence to Illocutionary Acts: some observations on the new look in Psycholinguistics. Paper presented to Canadian Psychological Association 38th Annual Meeting, Vancouver, B.C., June 9, 1977.

McCaffrey, A. Testing a Model of Communicative Competence in the Classroom. Paper presented to the Eastern Educational Research Association Annual Conference, Kiawah Island, South Carolina, Feb. 20-24, 1979.

McNeill, D. The acquisition of language. Harper & Row: N.Y. 1970.

Moffett, J. Teaching the Universe of Discourse. Houghton-Mifflin: Boston, 1968.

Moore, D.R. A comparison of three methods of teaching specific language skills to lower-class preschool children. Doctoral dissertation, Harvard Graduate School of Education, 1971.

Neisser, U. Cognitive Psychology. Appleton Century Crofts: N.Y., 1967.

O'Donnell, R.C., Griffin, W., & Norris, R.C. Syntax of kindergarten and elementary schoolchildren: a transformational analysis. National Council of Teachers of English, Monograph #8. NCTE: Champaign, Ill., 1967.

Palermo, D., & Molfese, D. Language acquisition from age five onward. Psychol. Bulletin, 18, 6, 1972.

Piaget, J. The language and thought of the child. Harcourt Brace: N.Y., 1926.

Piaget, J., & Inhelder, B. The child's conception of space. Routledge & Kegan Paul: London, 1956.

Pinard, A., & Laurendeau, M. The development of the concept of space in the child. International U. Press: N.Y., 1970.

Pinard, A., & Laurendeau, M. La pensée causale. Presses Universitaires:Paris, 1962.

Pipho, C. State Activity: Minimal Competency Testing. Report of the Dept. of Research and Information, Education Commission of the States: Denver, Colorado, July 1, 1979.

Rommetveit, R. Words, meanings and messages. Academic Press: N.Y., 1968.

Ruth, D.J. Language, intelligence and social class: a study of communication effectiveness within same-class and cross-class pairs. Unpublished paper, Harvard University, April, 1966.

Sapir, E. Language. Harcourt Brace: N.Y., 1921.

Searle, J. Speech Acts. Cambridge U. Press: London, 1969.

Selman, R. The growth of interpersonal understanding. Academic Press: N.Y., 1980.

Shantz, C., & Wilson, K. Training communication skills in young children. Child Development, 43, 1972.

Sharp, D., Cole, M., & Lave, C. Education and Cognitive Development: the evidence from Experimental Research. Monograph of the Society for Research in Child Development, #178, 44, 1-2. U. of Chicago Press: Chicago, 1979.

Shatz, M. On the development of communicative understanding-- an early strategy for interpreting and responding to messages. Cognitive Psych., 10,3, July, 1978.

Shatz, M., & Gelman, R. The development of communication skills: modifications in the speech of young children as function of listener. Monograph of the Society for Research in Child Development, #152, 38, 5, 1973.

Sinclair, A. Quelques aspects de la communication verbale chez les enfants de 4 ans. Paper presented to the Third International Child Language Symposium, U. of London, England, Sept., 1975.

Sinclair, H. Talking and Thinking. Lectures presented to the Division of Study and Research in Education, M.I.T., April 3-10, 1980.

Slobin, D. Questions of language development in cross-cultural perspective. Paper presented at the symposium on Language Learning in Cross-Cultural Perspective, Michigan State U., Sept. 25, 1968.

Spekman, N.J. The dyadic verbal communication skills of learning-disabled and normally achieving 4th & 5th grade boys. Paper presented at the Fifth Annual Conference on Language Development, Boston U., Oct., 1980.

Tamir, L. Language Development: New Directions. Human Development, 22, 1979.

Tough, J. The development of meaning: a study of children's use of language. Halsted Press: N.Y., 1977.

de Villiers, P.A. & J.C. Language Acquisition. Harvard U. Press:Cambridge, 1978.

Addenda:

Baldwin, C.P. & A.L. Interaction in the preschool years. Project Literacy Reports, Cornell U., Sept. 1966.

Bloom, L., Hood, L. & Lightbown, P. Imitation in language development: if, when, and why. Cognitive Psych., 6, 1974.

Bloom, L., Lightbown, P. & Hood, L. Structure and variation in child language. SRCD Monograph #160, 40, 2, 1975.

Fry, C.L. Training children to communicate to listeners. Child Devel., 37, 1966.

Fry, C.L. Training children to communicate to listeners with varying requirements. J. Genetic Psych., 114, 1969.

Hudson, L.M., Peyton, E. & Meisels, S.B. Role-taking as a determinant of prosocial behavior. Unpublished paper, Dept. of Ed. Psych., U. of Toledo, Ohio, 1978.

TABLE 1-a

COMPSTAT COMPOSITION OF TEST DYADS FOR BOTH LANGUAGE DOMAINS IN THE PRETEST, FOR THREE SCHOOLS

1. SEL

	COMPSTAT							
SCHOOLS	Lo/Lo	Lo/Mod	Lo/Hi	Mod/Hi	Mod/Mod	Hi/Hi	# Pairs	N*
A.	2	10	3	3	2	1	21	34
B.	3	8	1	5	7	0	24	28
C.	5	6	2	7	4	0	24	35
Total # Pairs	10	24	6	15	13	1	69	97 Ss
%	14	35	9	22	19	1		

2. LL

	COMPSTAT							
SCHOOLS	Lo/Lo	Lo/Mod	Lo/Hi	Mod/Hi	Mod/Mod	Hi/Hi	# Pairs	N*
A.	2	11	2	2	1	3	21	34
B.	2	10	0	7	5	0	24	28
C.	5	7	1	4	5	2	24	35
Total # Pairs	9	28	3	13	11	5	69	97 Ss
%	13	40	6	17	16	7		

(* = this is the actual number of Experimental and Control Ss, minus "SPARES", so the N is less than total number of Pairs.)

TABLE 1-b

COMPSTAT COMPOSITION OF TEST DYADS FOR BOTH LANGUAGE DOMAINS IN THE POST-TEST, FOR THREE SCHOOLS

1. STL

SCHOOLS	COMPSTAT							
	Lo/Lo	Lo/Mod	Lo/Hi	Mod/Hi	Mod/Mod	Hi/Hi	# Pairs	N*
A.	3	6	6	2	1	0	18	34
B.	3	4	3	2	1	1	14	28
C.	4	5	5	2	1	1	18	35
Total # Pairs	10	15	14	6	3	2	50	97 Ss
%	20	30	28	12	6	4		

2. LL

SCHOOLS	COMPSTAT							
	Lo/Lo	Lo/Mod	Lo/Hi	Mod/Hi	Mod/Mod	Hi/Hi	# Pairs	N*
A.	5	4	4	4	2	0	19	34
B.	2	7	2	5	3	0	19	28
C.	2	9	6	1	1	1	20	35
Total # Pairs	9	20	12	10	6	1	58	97 Ss
%	16	34	21	17	10	2		

(* = this is the actual number of Experimental and Control Ss, minus "SPARES", so the N is less than total number of Pairs.)

TABLE 1-c

Breakdown of Test Dyad Pairings by EXPGROUP and COMPSTAT for STL in the POST-TEST

1. Number of Pairings involving Low Ss:	with Lows	with Mods	with Highs	Total # Pairings
	10 (26%)	15 (38%)	14 (36%)	39

2. EXPGROUPS PAIRED WITH:

COMPSTAT:	LOW				MOD				HIGH				TOTALS				
STATUS:	E	C	£	(%)	E	C	£	(%)	E	C	£	(%)	E	%	C	%	N
(E=Exptl. C= Control)																	
EXPGROUPS																	
I	3	3	6	46	2	1	3	23	1	3	4	31	6	46	7	54	13
II	3	1	4	33	0	4	4	33	3	1	4	33	6	50	6	50	12
III	2	2	4	36	1	3	4	36	0	3	3	27	3	27	8	73	11
IV	6	0	6	46	3	1	4	31	1	2	3	23	10	77	3	23	13
Total	14	6	20	(41%)	6	9	15	(31%)	5	9	14	(28%)	25		24		49 Ss

TABLE 1-d

Breakdown of Test Dyad Pairings by EXPGROUPS and COMPSTAT for LL in the Post-TEST

1. Number of Pairings involving Low *Ss*:	with Lows	with Mods	with Highs	Total # Pairings
	9 (23%)	19 (47%)	12 (30%)	40

2. **EXPGROUPS PAIRED WITH:**

COMPSTAT:	LOW				MOD					HIGH				TOTALS						
STATUS:	E	C	Σ	(%)	E	C	S	Σ	(%)	E	C	Σ	(%)	E	(%)	C	(%)	S	(%)	N
(E=Exptl. C= Control S=SPARE)																				
EXPGROUPS																				
I	2	3	5	(38.5)	0	0	3	3	(23)	2	3	5	(38.5)	4	(31)	6	(46)	3	(23)	13
II	1	2	3	(25)	1	3	1	5	(42)	0	4	4	(33)	2	(17)	9	(75)	1	(8)	12
III	3	1	4	(36)	1	1	4	6	(55)	0	1	1	(9)	4	(36)	3	(27)	4	(36)	11
IV	6	0	6	(46)	2	1	2	5	(39)	2	0	2	(15)	10	(77)	1	(8)	2	(15)	13
Total	12	6	18	(37)	4	5	10	19	(39)	4	8	12	(24)	20	(41)	19	(39)	10	(20)	49 Ss

TABLE 2

PERCEPTION: ACCURACY x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\overline{X}$	S.D.	$\overline{X}$	S.D.
Lo/Lo	I	13	1.92	1.04	2.23	0.93
Lo/Mod	II	12	2.58	1.16	2.83	1.11
Lo/Hi	III	11	2.18	1.33	2.55	1.51
Lo/Control	IV	13	1.92	0.64	2.46	1.33
	Total	49	2.14	1.06	2.51	1.21
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	1.072	F =	0.511
			p =	0.371	p =	0.677
Position			F =		F =	
			p =		p =	

TABLE 3

PERCEPTION: SPEED x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	11.04	6.91	9.75	3.79
Lo/Mod	II	12	9.16	3.76	10.39	3.99
Lo/Hi	III	11	9.67	5.40	9.54	3.93
Lo/Control	IV	13	10.98	5.44	10.87	5.39
	Total	49	10.25	5.42	10.16	4.24
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	0.358	F =	0.242
			p =	0.783	p =	0.867
Position			F =		F =	
			p =		p =	

TABLE 4

PERCEPTION: SPEED x ACCURACY x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	8.66	4.69	9.55	4.97
Lo/Mod	II	12	10.06	5.36	10.52	4.01
Lo/Hi	III	11	7.45	4.41	9.29	3.48
Lo/Control	IV	13	10.21	6.69	11.01	5.47
	Total	49	9.14	5.35	10.12	4.51
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	0.679	F =	0.378
			p =	0.569	p =	0.770
Position			F =		F =	
			p =		p =	

TABLE 5

PERCEPTION: SPEED x INACCURACY x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	11.44	7.95	10.08	4.33
Lo/Mod	II	12	7.69	3.79	10.26	6.46
Lo/Hi	III	11	13.01	14.01	7.66	5.70
Lo/Control	IV	13	10.82	5.50	8.96	3.74
	Total	49	10.71	8.43	9.28	5.05
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	0.806	F =	0.639
			p =	0.497	p =	0.594
Position			F =		F =	
			p =		p =	

TABLE 6

PERCEPTION: COMPOSITE ACCURACY x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	13.66	6.76	15.69	5.58
Lo/Mod	II	12	17.63	6.56	19.10	6.15
Lo/Hi	III	11	15.02	8.28	17.12	8.35
Lo/Control	IV	13	13.96	4.04	16.76	7.63
	Total	49	15.02	6.49	17.13	6.86
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	0.949	F =	0.516
			p =	0.425	p =	0.674
Position			F =		F =	
			p =		p =	

TABLE 7

PERCEPTION: INTERCORRELATIONS AMONG MEASURES, PRETEST AND POSTTEST (N = 49)

	1	2	3	4	5	6	7	8	9	10
1. ACCURACY/PRE	--	.27	.17	.16	.37	.11	.11	.09	.99	.26
2. ACCURACY/POST		--	.28	.25	.17	.20	.32	-.20	.25	.99
3. SPEED/PRE			--	.27	.75	.15	.81	.38	.17	.29
4. SPEED/POST				--	.22	.91	.19	.43	.15	.25
5. SPEED X ACCURACY/PRE					--	.10	.45	.32	.38	.18
6. SPEED X ACCURACY/POST						--	.14	.25	.10	.20
7. SPEED X INACCURACY/PRE							--	.04	.11	.31
8. SPEED X INACCURACY/POST								--	.07	-.16
9. COMPOSITE ACCURACY/PRE									--	.25
10. COMPOSITE ACCURACY/POST										--

(Note: When r >.24, p <.05.)

TABLE 8

COGNITION: HOUSE CONSTRUCTION TASK (max.=3) x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	1.23	0.44	1.38	0.65
Lo/Mod	II	12	1.25	0.62	2.00	0.85
Lo/Hi	III	11	1.64	0.81	2.09	1.04
Lo/Control	IV	13	1.31	0.63	1.46	0.88
	Total	49	1.35	0.63	1.71	0.89
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F = 1.032 (3,45)		F = 2.160 (3,45)	
			p = 0.388		p = 0.106	
Position			F =		F =	
			p =		p =	

TABLE 9

COGNITION: HOUSE PHOTOS TASK (max.=12) x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	**$\bar{X}$**	**S.D.**	**$\bar{X}$**	**S.D.**
Lo/Lo	I	13	8.38	2.06	8.69	1.11
Lo/Mod	II	12	8.00	1.04	9.08	2.07
Lo/Hi	III	11	8.82	1.89	9.64	2.16
Lo/Control	IV	13	8.62	2.06	8.92	1.98
	Total	49	8.45	1.79	9.06	1.86
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F = 0.434 (3,45)		F = 0.547 (3,45)	
			p = 0.730		p = 0.653	
Position			F =		F =	
			p =		p =	

TABLE 10

COGNITION: MOUNTAIN PERSPECTIVES TASK (max.=9) x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	4.62	1.39	4.69	1.25
Lo/Mod	II	12	4.67	1.44	6.00	1.81
Lo/Hi	III	11	5.18	1.47	6.55	1.86
Lo/Control	IV	13	5.08	1.85	4.92	1.85
	Total	49	4.88	1.55	5.49	1.70
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F = 0.410 (3,45)		F = 3.20 (3,45)	
			p = 0.746		p = 0.03	
Position			F =		F =	
			p =		p =	

TABLE 11

COGNITION: No. of 1-Scores (out of 8) x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	**$\overline{X}$**	**S.D.**	**$\overline{X}$**	**S.D.**
Lo/Lo	I	13	3.23	2.13	2.69	1.18
Lo/Mod	II	12	3.33	1.07	2.00	1.595
Lo/Hi	III	11	2.27	1.56	2.00	2.28
Lo/Control	IV	13	2.23	1.69	3.00	1.68
	Total	49	2.78	1.62	2.45	1.70
	Range					
ANOVAS						
Main			F = p =		F = p =	
Expgroup			F = 1.562 (3.45) p = .25		F = 1.073 p = .370	
Position			F = p =		F = p =	

TABLE 12

COGNITION: No. of 2-Scores (out of 8) x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	3.39	1.71	3.85	0.987
Lo/Mod	II	12	3.42	0.669	2.92	1.31
Lo/Hi	III	11	3.82	1.40	2.91	2.12
Lo/Control	IV	13	3.92	1.71	2.69	1.25
	Total	49	3.63	1.40	3.10	1.45
	Range					
ANOVAS						
Main			F =		F =	
			p =		p =	
Expgroup			F =	0.450 (3,45)	F =	1.623
			p =	.50	p =	0.197
Position			F =		F =	
			p =		p =	

TABLE 13

COGNITION: No. of 3-Scores (out of 8) x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	1.38	1.19	1.54	1.05
Lo/Mod	II	12	1.25	0.866	3.08	2.31
Lo/Hi	III	11	1.91	2.07	3.64	2.69
Lo/Control	IV	13	1.85	1.41	2.31	1.38
	Total	49	1.59	1.38	2.59	1.93
	Range					
ANOVAS						
Main			F = p =		F = p =	
Expgroup			F = 0.638 (3,45) p = 0.50		F = 2.721 (3,45) p = .055	
Position			F = p =		F = p =	

TABLE 14

COGNITION: SUMMARY MEASURE (COGSUM) x EXPGROUPS.
(max.=24)

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	14.23	3.004	14.77	2.09
Lo/Mod	II	12	13.92	1.83	17.08	3.75
Lo/Hi	III	11	15.64	3.38	18.27	3.82
Lo/Control	IV	13	15.00	2.92	15.31	2.81
	Total	49	14.67	2.75	16.26	3.16
	Range					
ANOVAS						
Main			F = p =		F = p =	
Expgroup			F = 0.871 (3,45) p = 0.50		F = 3.125 (3,45) p = 0.035	
Position			F = p =		F = p =	

TABLE 15

COGNITION: INTERCORRELATIONS AMONG MEASURES, PRETEST and POST-TEST (N=49)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. HOUSE CONSTRUCTION TASK/PRETEST	--	.26	.26	.12	.15	.17	-.39	-.23	.05	-.18	.40	.26	.47	.24
2. " /POST-TEST		--	.39	.37	-.06	.07	-.23	-.31	-.09	-.24	.36	.53	.28	.52
3. HOUSE PHOTOS TASK/PRETEST			--	.43	.18	.06	-.59	-.38	-.11	-.13	.80	.35	.80	.39
4. " /POST-TEST				--	.14	.24	-.29	-.61	-.07	-.31	.42	.78	.38	.82
5. MTN. PERSPECTIVES TASK/PRETEST					--	.43	-.69	-.32	.34	-.19	.46	.32	.69	.30
6. " /POST-TEST						--	-.25	-.58	.01	-.11	.27	.60	.31	.70
7. 1-Score/PRE							--	.46	-.59	.04	-.57	-.30	-.85	-.37
8. " /POST								--	-.10	-.18	-.43	-.55	-.47	-.76
9. 2-Score/PRE									--	.26	-.33	-.16	.13	-.06
10. " /POST										--	-.31	-.63	-.23	-.31
11. 3-Score/PRE											--	.51	.86	.49
12. " /POST												--	.46	.93
13. COGSUM/PRE													--	.47
14. " /POST														--

TABLE 16

Summary Raw Data for the Basic Proposition Scores in the STL Domain for EXPGROUPS (N=49), both Pretest and Post-test

	SCORE TOTALS							
	+	%	0	%	-	%	Total	N
PRETEST								
EXPGROUPS								
I	172	(58)	94	(32)	29	(10)	295	13
II	141	(57)	81	(33)	26	(10)	248	12
III	148	(66)	59	(26)	16	(7)	223	11
IV	158	(69)	47	(21)	23	(10)	228	13
Total	619	(62)	281	(28)	94	(9)	994	49 *Ss*
POST-TEST								
EXPGROUPS								
I	244	(75)	75	(23)	5	(2)	324	13
II	202	(70)	77	(27)	9	(3)	288	12
III	180	(72)	58	(23)	11	(4)	249	11
IV	207	(78)	49	(18)	10	(4)	266	13
Total	833	(74)	259	(23)	35	(3)	1127	49 *Ss*

TABLE 17

LANGUAGE: STL, ACCURACY OF PREPOSITIONAL USAGE x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	4.77	1.01	5.54	0.66
Lo/Mod	II	12	4.58	1.38	5.00	0.95
Lo/Hi	III	11	4.64	1.12	5.27	0.79
Lo/Control	IV	13	4.23	1.36	5.31	1.18
	Total	49	4.55	1.21	5.29	0.91
	Range		2 - 6		2 - 6	
ANOVAS						
Main			F = 0.311 (4,41)		F = 0.652 (4,41)	
			p = 0.869		p = 0.629	
Expgroup			F = 0.407 (3,41)		F = 0.585 (3,41)	
			p = 0.748		p = 0.628	
Position			F = 0.009 (1,41)		F = 0.540 (1,41)	
			p = 0.926		p = 0.467	

TABLE 18

LANGUAGE: STL, QUANTITY OF PROPOSITIONS x EXPGROUP.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	22.69	6.76	24.92	8.37
Lo/Mod	II	12	20.67	4.92	24.00	7.91
Lo/Hi	III	11	20.21	5.42	22.64	6.92
Lo/Control	IV	13	17.54	4.98	20.46	4.18
	Total	49	20.29	5.74	23.00	7.01
	Range		7 - 43		6 - 43	
ANOVAS						
Main			F = 1.31 (4,41) p = 0.282		F = 0.780 (4,41) p = 0.545	
Expgroup			F = 1.73 (3,41) p = 0.176		F = 0.950 (3,41) p = 0.425	
Position			F = 0.053 (1,41) p = 0.819		F = 0.338 (1,41) p = 0.564	

TABLE 19

LANGUAGE: STL, PROPOSITIONAL QUALITY x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	13.23	3.49	18.77	5.19
Lo/Mod	II	12	11.75	1.96	16.83	3.51
Lo/Hi	III	11	13.45	3.93	16.36	4.95
Lo/Control	IV	13	12.15	3.48	15.92	2.90
	Total	49	12.63	3.27	17.00	4.25
	Range		5 - 20		6 - 28	
ANOVAS						
Main			F = 0.534 (4,41)		F = 1.19	
			p = 0.712		p = 0.329	
Expgroup			F = 0.711 (3,41)		F = 0.985	
			p = 0.551		p = 0.409	
Position			F = 0.044 (1,41)		F = 1.47	
			p = 0.835		p = 0.232	

TABLE 20

LANGUAGE: STL, MESSAGE EFFICIENCY I(%) x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	59.94	11.58	76.96	12.45
Lo/Mod	II	12	58.02	7.29	72.96	13.09
Lo/Hi	III	11	68.17	15.74	73.57	16.61
Lo/Control	IV	13	69.67	12.26	79.29	14.65
	Total	49	63.89	12.69	75.84	14.00
	Range		29 - 85%		44 - 100%	
ANOVAS						
Main			F = 2.014 (4,41) p = 0.110		F = 0.524 (4,41) p = 0.719	
Expgroup			F = 2.66 (3,41) p = 0.061		F = 0.375 (3,41) p = 0.772	
Position			F = 0.002 (1,41) p = 0.963		F = 0.512 (1,41) p = 0.478	

TABLE 21

LANGUAGE: GAIN SCORES (%) for Efficiency II Measure in the STL Domain, for EXPGROUPS (N=49)

COMPSTAT	EXPGROUP	N	Mean Gain	S.D.
Lo/Lo	I	13	17.02	12.56
Lo/Mod	II	12	14.95	12.69
Lo/Hi	III	11	5.41	17.93
Lo/Control	IV	13	9.62	14.56
	Total	49	11.94%	14.72

One-way ANOVA:

	D.F.	Mean Square	F	p.
Between Groups	3	327.45	1.56	$< .21$
Within Groups	45	209.43		

TABLE 22

LANGUAGE: STL, MESSAGE EFFICIENCY III x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	20.29	3.24	24.68	3.11
Lo/Mod	II	12	19.83	1.99	23.68	3.36
Lo/Hi	III	11	22.38	4.38	23.77	4.09
Lo/Control	IV	13	22.86	3.08	25.21	3.73
	Total	49	21.33	3.41	24.37	3.52
	Range					
ANOVAS						
Main			F = 1.84 (4,41) p = 0.140		F = 0.505 (4,41) p = 0.733	
Expgroup			F = 2.43 (3,41) p = 0.079		F = 0.366 (3,41) p = 0.778	
Position			F = 0.001 (1,41) p = 0.970		F = 0.467 (1,41) p = 0.498	

TABLE 23

LANGUAGE: STL, PROPOSITIONAL AMBIGUITY x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	7.23	4.87	5.77	4.83
Lo/Mod	II	12	6.75	3.65	6.42	6.29
Lo/Hi	III	11	5.36	4.63	5.27	4.69
Lo/Control	IV	13	3.62	2.14	3.77	3.92
	Total	49	5.73	4.09	5.29	4.93
	Range		2 - 16		0 - 24	
ANOVAS						
Main			F = 1.71 (4,41) p = 0.166		F = 0.466 p = 0.760	
Expgroup			F = 2.04 (3,41) p = 0.123		F = 0.617 p = 0.608	
Position			F = 0.689 (1,41) p = 0.411		F = 0.001 p = 0.974	

TABLE 24

LANGUAGE: STL, MESSAGE INEFFICIENCY I x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	0.301	0.118	0.210	0.123
Lo/Mod	II	12	0.313	0.114	0.234	0.145
Lo/Hi	III	11	0.243	0.177	0.214	0.155
Lo/Control	IV	13	0.205	0.091	0.171	0.148
	Total	49	0.265	0.131	0.206	0.140
	Range					
ANOVAS						
Main			F = 1.725 (4,41) p = 0.162		F = 0.359 (4,41) p = 0.836	
Expgroup			F = 1.806 (3,41) p = 0.149		F = 0.372 (3,41) p = 0.774	
Position			F = 1.142 (1,41) p = 0.292		F = 0.159 p = 0.692	

TABLE 25

LANGUAGE: STL, MESSAGE INEFFICIENCY II x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	11.26	3.80	8.08	4.33
Lo/Mod	II	12	11.67	3.81	8.85	5.08
Lo/Hi	III	11	9.07	5.80	8.09	5.63
Lo/Control	IV	13	7.97	3.33	6.53	5.39
	Total	49	9.996	4.38	7.86	5.02
	Range					
ANOVAS						
Main			F = 1.88 (4,41) p = 0.132		F = 0.399 (4,41) p = 0.809	
Expgroup			F = 2.03 (3,41) p = 0.125		F = 0.392 (3,41) p = 0.759	
Position			F = 1.21 (1,41) p = 0.279		F = 0.241 (1,41) p = 0.626	

TABLE 26

LANGUAGE: STL, MESSAGE ORGANISATION (max.=18) x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	11.31	3.45	15.77	2.38
Lo/Mod	II	12	10.17	2.69	14.83	2.04
Lo/Hi	III	11	11.46	3.39	12.82	4.24
Lo/Control	IV	13	11.00	3.34	14.46	3.15
	Total	49	10.98	3.17	14.53	3.12
	Range		4 - 18		4 - 18	
ANOVAS						
Main			F = 0.265 (4,41) p = 0.899		F = 1.44 (4,41) p = 0.239	
Expgroup			F = 0.347 (3,41) p = 0.791		F = 1.62 (3,41) p = 0.200	
Position			F = 0.002 (1,41) p = 0.964		F = 0.203 (1,41) p = 0.655	

TABLE 27

LANGUAGE: STL, COMPOSITE COMPETENCE (%) x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	60.37	11.39	80.81	9.89
Lo/Mod	II	12	56.79	7.84	75.98	7.90
Lo/Hi	III	11	65.09	14.07	70.99	17.49
Lo/Control	IV	13	64.51	14.52	79.75	14.01
	Total	49	61.65	12.32	77.14	12.89
	Range		37 - 90		35 - 95	
ANOVAS						
Main			F = 0.843 (4,41) p = 0.506		F = 1.15 p = 0.349	
Expgroup			F = 1.11 (3,41) p = 0.357		F = 1.02 p = 0.394	
Position			F = 0.002 (1,41) p = 0.965		F = 0.509 p = 0.480	

TABLE 28

LANGUAGE: STL, COMPOSITE EFFICIENCY II x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	20.41	3.08	25.66	2.37
Lo/Mod	II	12	19.48	2.14	24.50	1.97
Lo/Hi	III	11	21.63	3.71	23.08	4.58
Lo/Control	IV	13	21.46	3.96	25.34	3.45
	Total	49	20.73	3.31	24.71	3.26
	Range					
ANOVAS						
Main			F = 0.76 (4,41) p = .56		F = 1.19 (4,41) p = .33	
Expgroup			F = 1.003 (3,41) p = .40		F = 1.08 (3,41) p = .37	
Position			F = 0.002 (1,41) p = .97		F = 0.503 (1,41) p = .48	

TABLE 29

LANGUAGE: COMPOSITE COMPETENCE GAIN SCORES for STL Domain, FOR EXPGROUPS (N=49)

COMPSTAT	EXPGROUP	N	Mean Gain	S.D.
Lo/Lo	I	13	20.45	10.42
Lo/Mod	II	12	19.19	10.45
Lo/Hi	III	11	5.91	15.05
Lo/Control	IV	13	15.24	15.62
	Total	49	15.49	13.84

One-way ANOVA:

	D.F.	Mean Square	F	p.
Between Groups	3	498.23	2.91	$<.04$
Within Groups	45	171.05		

TABLE 30

LANGUAGE: STL, PREDICTED QUALITY OF PROPOSITIONS x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	-1.15	2.65	1.30	2.86
Lo/Mod	II	12	-1.82	0.98	-0.27	2.70
Lo/Hi	III	11	+0.05	3.24	-0.19	3.60
Lo/Control	IV	13	-0.16	2.28	+0.24	2.88
	Total	49	-0.78	2.46	+0.30	2.98
	Range					
ANOVAS						
Main			F = 1.09 (4,41) p = 0.374		F = 0.877 p = 0.486	
Expgroup			F = 1.45 (3,41) p = 0.244		F = 0.501 p = 0.684	
Position			F = 0.005 (1,41) p = 0.943		F = 1.36 p = 0.250	

TABLE 31

SUMMARY RAW DATA for Positive and Ambiguous Propositions in the LL Domain, by TASKS, PRETEST and POST-TEST, for EXPGROUPS (N=49)

PRETEST

	# Propositions				
TASKS	**Positive**	**%**	**Ambiguous**	**%**	**Total**
1. Prod. Bd.	153	(35)	283	(65)	436 (27% of Domain Total
2. Picture Boxes	50	(8)	598	(92)	648 (41% of Domain Total
3. Cartoon Boxes	166	(32)	350	(68)	516 (32% of Domain Total
TOTALS FOR LL DOMAIN	369	(23)	1231	(77)	1600

POST-TEST

TASKS	**Positive**	**%**	**Ambiguous**	**%**	**Total**
1. Prod. Bd.	193	(39)	307	(61)	500 (30% of Domain Total
2. Picture Boxes	132	(22)	478	(78)	610 (37% of Domain Total
3. Cartoon Boxes	228	(42)	309	(58)	537 (33% of Domain Total
TOTALS FOR LL DOMAIN	553	(34)	1094	(66)	1647

TABLE 32

LANGUAGE: LL, QUANTITY OF PROPOSITIONS x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	36.62	13.81	36.92	8.00
Lo/Mod	II	12	32.83	16.39	32.92	7.74
Lo/Hi	III	11	31.91	9.96	34.82	8.21
Lo/Control	IV	13	29.15	11.39	30.15	6.99
	Total	49	32.65	13.04	33.67	7.91
	Range		13 - 75		17 - 52	
ANOVAS						
Main			F = 1.15 (4,41) p = 0.349		F = 1.496 (4,41) p = 0.221	
Expgroup			F = 0.789 (3,41) p = 0.507		F = 1.21 (3,41) p = 0.319	
Position			F = 2.49 p = 0.122		F = 0.899 (1,41) p = 0.349	

TABLE 33

LANGUAGE: LL, DISCRIMINABILITY INDEX x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	24.23	4.09	28.62	2.26
Lo/Mod	II	12	24.67	3.49	29.0	1.81
Lo/Hi	III	11	22.36	3.47	29.0	1.34
Lo/Control	IV	13	21.85	5.31	26.31	3.25
	Total	49	23.29	4.24	28.18	2.52
	Range		8 - 31		20 - 31	
ANOVAS						
Main			F = 2.89 (4,41)		F = 2.84 (4,41)	
			p = 0.034		p = 0.036	
Expgroup			F = 1.18 (3,41)		F = 3.43 (3,41)	
			p = 0.331		p = 0.026	
Position			F = 7.23 (1,41)		F = 0.156 (1,41)	
			p = 0.01		p = 0.695	

TABLE 34

LANGUAGE: LL, PROPOSITIONAL QUALITY x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	7.77	3.24	11.92	4.27
Lo/Mod	II	12	7.58	2.87	11.67	4.54
Lo/Hi	III	11	8.09	3.91	12.09	3.33
Lo/Control	IV	13	6.62	3.55	9.31	2.75
	Total	49	7.49	3.34	11.20	3.85
	Range					
ANOVAS						
Main			F = 2.44 (4,41)		F = 1.18 (4,41)	
			p = 0.062		p = 0.332	
Expgroup			F = 0.887 (3,41)		F = 1.58 (3,41)	
			p = 0.456		p = 0.209	
Position			F = 8.31 (1,41)		F = 0.238 (1,41)	
			p = 0.006		p = 0.629	

TABLE 35

LANGUAGE: LL, MESSAGE EFFICIENCY x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	1.99	0.883	3.01	0.818
Lo/Mod	II	12	2.35	0.759	3.09	0.725
Lo/Hi	III	11	2.24	0.800	2.97	0.834
Lo/Control	IV	13	1.90	0.923	2.91	1.04
	Total	49	2.11	0.841	2.99	0.842
	Range		0 - 3.875		1.1 - 5.08	
ANOVAS						
Main			F = 0.534 (4,41) p = 0.711		F = 0.424 (4,41) p = 0.790	
Expgroup			F = 0.682 (3,41) p = 0.568		F = 0.148 (3,41) p = 0.930	
Position			F = 0.063 (1,41) p = 0.803		F = 1.41 (1,41) p = 0.242	

TABLE 36

LANGUAGE: LL, MESSAGE INEFFICIENCY x EXPGROUPS.

			PRE		POST	
COMPSTAT	**EXPGROUP**	**N**	$\bar{X}$	**S.D.**	$\bar{X}$	**S.D.**
Lo/Lo	I	13	6.01	0.883	4.99	0.818
Lo/Mod	II	12	5.65	0.760	4.91	0.725
Lo/Hi	III	11	5.58	0.671	5.03	0.834
Lo/Control	IV	13	6.098	0.923	5.04	1.06
	Total	49	5.85	0.827	4.99	0.847
	Range		4.125 - 8.0		2.92 - 6.9	
ANCOVAS						
Main			F = 0.816 (4,41)		F = 0.323 (4,41)	
			p = 0.523		p = 0.861	
Expgroup			F = 1.09 (3,41)		F = 0.077 (3,41)	
			p = 0.366		p = 0.972	
Position			F = 0.016 (1,41)		F = 1.11 (1,41)	
			p = 0.899		p = 0.297	

TABLE 37

LANGUAGE: LL, MESSAGE ORGANISATION x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	7.71	2.45	4.69	1.99
Lo/Mod	II	12	7.31	1.87	4.25	1.82
Lo/Hi	III	11	6.98	2.27	4.81	1.54
Lo/Control	IV	13	8.56	2.99	5.91	1.69
	Total	49	7.67	2.44	4.95	1.85
	Range		2.38 - 15		1.38 - 8.75	
ANOVAS						
Main			F = 0.727 (4,41) p = 0.579		F = 2.18 (4,41) p = 0.088	
Expgroup			F = 0.921 (3,41) p = 0.439		F = 2.55 (3,41) p = 0.069	
Position			F = 0.262 (1,41) p = 0.611		F = 2.05 (1,41) p = 0.159	

TABLE 38

LANGUAGE: LL, PREDICTED QUALITY OF PROPOSITIONS x EXPGROUPS.

			PRE		POST	
COMPSTAT	EXPGROUP	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Lo/Lo	I	13	-1.85	3.11	1.24	3.06
Lo/Mod	II	12	-1.28	2.46	1.78	3.78
Lo/Hi	III	11	-0.59	2.99	1.83	3.10
Lo/Control	IV	13	-1.52	2.44	-0.02	2.58
	Total	49	-1.34	2.71	1.17	3.14
	Range					
ANOVAS						
Main			F = 1.18 (4,41) p = 0.333		F = 1.01 (4,41) p = 0.416	
Expgroup			F = 0.635 (3,41) p = 0.596		F = 1.17 (3,41) p = 0.333	
Position			F = 3.42 (1,41) p = 0.072		F = 1.16 (1,41) p = 0.288	

TABLE 39 (PRETEST)

PERCEPTION: One-way ANOVAS for Accuracy and Speed, by COMPSTAT

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
	49	Low	2.14	1.06	F	2.39		
1. ACCURACY (TPCB1)	24	Mod	2.63	1.09	df	2,94		
	24	High	2.67	1.31	p <	.097		
	97	Total	2.39	1.13				
VARIABLE								
		Low	10.25	5.42	F	3.88		
2. SPEED (secs.) (MTPSB)		Mod	10.94	4.48	df	2,94		
		High	14.09	6.84	p <	.02		
		Total	11.37	5.60				
VARIABLE								
		Low	9.14	5.35	F	3.14		
3. SPEED X ACCURACY (MTPSB1)		Mod	10.82	5.09	df	2,94		
		High	12.82	7.67	p <	.047		
		Total	10.47	5.95				

TABLE 39 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
4. SPEED X INACCURACY (MTPSB2)	49	Low	10.71	8.43	F	1.51		
	24	Mod	10.23	5.03	df	2,94		
	24	High	13.78	9.35	p <	.23		
	97	Total	11.35	8.04				
VARIABLE								
5. COMPOSITE ACCURACY RATIO (PACCB)		Low	15.01	6.49	F	2.16		
		Mod	17.89	6.28	df	2,94		
		High	17.90	7.93	p <	.12		
		Total	16.44	6.91				
VARIABLE								
		Low			F			
		Mod			df			
		High			p			
		Total						

TABLE 40 (PRETEST)

COGNITION: One-way ANOVAS for TASKS, SCORES and SUMMARY MEASURE (COGSUM), by COMPSTAT

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
1. HOUSE CONSTRUCTION TASK (max. poss. = 3) (HCONCOG)	49	Low	1.35	0.63	F	3.15		
	24	Mod	1.50	0.78	df	2,94		
	24	High	1.79	0.78	p <	.046		
	97	Total	1.49	0.71				
VARIABLE								
2. HOUSE PHOTOS TASK (max poss. = 12) (HPICCOG)		Low	8.45	1.79	F	5.82		
		Mod	9.13	1.15	df	2,94		
		High	9.79	1.59	p <	.004		
		Total	8.95	1.61				
VARIABLE								
3. MOUNTAIN PERSPECTIVES TASK (max. poss. = 9) (MTPCOG)		Low	4.88	1.52	F	5.36		
		Mod	4.79	1.47	df	2,94		
		High	6.04	1.68	p <	.006		
		Total	5.14	1.55				

TABLE 40 (cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
4. 1-score max. = 8 (S1COG)	49	Low	2.78	1.69	F	3.53		
	24	Mod	2.54	1.38	df	2,94		
	24	High	1.71	1.68	p <	.03		
	97	Total	2.45	1.61				
VARIABLE								
5. 2-score max. = 8 (S2COG)		Low	3.63	1.42	F	2.47		
		Mod	3.58	1.06	df	2,94		
		High	2.92	1.41	p <	.08		
		Total	3.44	1.36				
VARIABLE								
6. 3-score max. = 8 (S3COG)		Low	1.59	1.41	F	12.25		
		Mod	1.92	1.32	df	2,94		
		High	3.38	1.69	p <	.001		
		Total	2.11	1.63				
VARIABLE								
7. COGSUM (max. poss. = 24)		Low	14.67	2.82	F	8.95		
		Mod	15.42	2.52	df	2,94		
		High	17.63	3.05	p <	.001		
		Total	15.59	2.78				

TABLE 41 (PRETEST)

LANGUAGE: STL (ACTION-SEQUENCE), Summary Raw Data for Positive/Ambiguous/Negative Propositions, by COMPSTAT (N = 97)

COMPSTAT	Mean Number PROPOSITIONS			
	+	0	-	Total
Low (49)	12.63	5.73	1.92	20.29
Mod (24)	15.21	4.13	1.46	20.79
High (24)	18.25	2.5	1.08	21.83
Total (97)	14.66	4.54	1.6	20.8

TABLE 42 (PRETEST)

LANGUAGE: STL (ACTION-SEQUENCE), Two-way ANOVAs

by COMPSTAT and Test Position

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
1. Prepositional Usage (max. = 6) (SPUPREP)	49	Low	4.55	1.21	F 5.34	7.36	0.18	.274
	24	Mod	5.17	0.64	df 3,91	2,91	1,91	2,91
	24	High	5.46	0.59	p<.002	.001	.999	.999
	97	Total	4.93	1.03				
VARIABLE								
2. Quantity (X1)		Low	20.29	5.74	F .614	.604	.537	.598
		Mod	20.79	5.46	df 3,91	2,91	1,91	2,91
		High	21.83	4.54	p< .999	.999	.999	.999
		Total	20.79	5.38				
VARIABLE								
3. Quality (SPUPOS)		Low	12.63	3.27	F 14.96	22.23	.039	.018
		Mod	15.21	2.96	df 3,91	2,91	1,91	2,91
		High	18.25	3.83	p< .001	.001	.999	.999
		Total	14.66	4.04				

TABLE 42 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
4. EFFICIENCY I (PURATIO)	49	Low	63.89	12.69	F 14.37	21.54	.450	.653
	24	Mod	75.70	14.99	df 3,91	2,91	1,91	2,91
	24	High	84.20	10.60	p < .001	.001	.999	.999
	97	Total	71.84	15.24				
VARIABLE								
5. EFFICIENCY III (STLQUAL)		Low	21.33	3.41	F 13.67	20.48	.473	.718
		Mod	24.32	3.84	df 3,91	2,91	1,91	2,91
		High	26.46	2.5	p < .001	.001	.493	.490
		Total	23.34	3.95				
VARIABLE								
6. AMBIGUITY (SPUZER)		Low	5.73	4.09	F 4.64	6.47	1.55	.426
		Mod	4.13	4.18	df 3,91	2,91	1,91	2,91
		High	2.50	2.19	p < .005	.003	.214	.999
		Total	4.54	3.94				

TABLE 42 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
7. INEFFICIENCY I (STLAMBA)	49	Low	0.27	.13	F 9.79	14.24	1.84	.286
	24	Mod	0.18	.14	df 3,91	2,91	1,91	2,91
	24	High	0.11	.09	p < .001	.001	.174	.999
	97	Total	0.20	.14				
VARIABLE								
8. INEFFICIENCY II (STLAMB1)		Low	9.99	4.38	F 10.43	15.22	1.78	.247
		Mod	6.78	4.99	df 3,91	2,91	1,91	2,91
		High	4.29	3.38	p < .001	.001	.185	.782
		Total	7.79	4.91				
VARIABLE								
9. ORGANISATION (FMSEQ)		Low	10.98	3.17	F 16.99	25.36	.001	.258
		Mod	13.79	2.67	df 3,91	2,91	1,91	2,91
		High	15.92	2.12	p < .001	.001	.999	.999
		Total	12.89	3.49				

TABLE 42 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
10. COMPOSITE COMPETENCE (PUPMEFFS)	49	Low	61.65	12.32	F 26.39	39.58	.273	.085
	24	Mod	75.30	10.05	df 3,91	2,91	1,91	2,91
	24	High	85.71	9.06	p < .001	.001	.603	.918
	97	Total	70.98	14.94				
VARIABLE								
11. COMPOSITE EFFICIENCY II (STLEFF2B)		Low	20.73	3.31	F 25.29	37.93	.262	.088
		Mod	24.31	2.52	df 3,91	2,91	1,91	2,91
		High	26.83	2.15	p < .001	.001	.610	.916
		Total	23.13	3.85				
VARIABLE								
12. PREDICTED QUALITY (PRSTLQL)		Low	-0.78	2.46	F 20.97	31.44	.130	.261
		Mod	1.59	2.64	df 3,91	2,91	1,91	2,91
		High	4.22	2.57	p < .001	.001	.719	.771
		Total	1.04	3.25				

TABLE 43 (PRETEST)

LANGUAGE: LL (Picture Description), Summary Raw Data for Positive and Ambiguous Propositions by COMPSTAT (N = 97)

COMPSTAT	Mean Number PROPOSITIONS		
	+	0	Total
Low (49)	7.49	25.12	32.65
Mod (24)	8.71	26.96	35.67
High (24)	10.71	28.17	38.88
Total (97)	8.59	26.35	34.94

TABLE 44 (PRETEST)

LANGUAGE: LL (Picture Description), Two-way ANOVAS

by COMPSTAT and Test Position

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS			
	N		$\bar{X}$	S.D.	MAIN	COMPSTAT	POSITION	INTERACTION
1. QUANTITY	49	Low	32.65	13.04	F 2.10	1.64	2.20	.452
(SSLLTOT)	24	Mod	35.67	11.27	df 3,91	2,91	1,91	2,91
	24	High	38.88	12.54	p < .104	.199	.137	.999
	97	Total	34.94	12.64				
VARIABLE								
		Low	23.29	4.24	F 17.02	18.89	7.68	2.44
2. DISCRIMINABILITY		Mod	26.29	2.22	df 3,91	2,91	1,91	2,91
INDEX		High	28.46	2.34	p < .001	.001	.007	.093
(SSVCDS)		Total	25.31	4.04				
VARIABLE								
		Low	7.49	3.34	F 6.93	7.49	3.59	3.52
3. QUALITY		Mod	8.71	2.84	df 3,91	2,91	1,91	2,91
(SSPOSVC)		High	10.71	3.41	p < .001	.001	.06	.034
		Total	8.59	3.47				

TABLE 44 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS				
	N		$\bar{X}$	S.D.		MAIN	COMPSTAT	POSITION	INTERACTION
4. EFFICIENCY (SSLLQAL)	49	Low	2.11	.84	F	1.86	2.78	.052	.245
	24	Mod	2.37	.71	df	3,91	2,91	1,91	2,91
	24	High	2.55	.62	p <	.143	.067	.821	.783
	97	Total	2.28	.78					
VARIABLE									
5. INEFFICIENCY (SSLLAMB)		Low	5.85	.83	F	1.40	2.09	.042	.06
		Mod	5.59	.71	df	3,91	2,91	1,91	2,91
		High	5.49	.58	p <	.25	.13	.838	.946
		Total	5.69	.75					
VARIABLE									
6. ORGANISATION (LLUL)		Low	7.67	2.44	F	10.69	16.03	.443	.962
		Mod	6.07	1.16	df	3,91	2,91	1,91	2,91
		High	5.02	1.41	p <	.001	.001	.507	.386
		Total	6.62	2.25					

TABLE 44 (Cont'd)

VARIABLE	SCORES x COMPSTAT				ANOVA EFFECTS				
	N		$\overline{X}$	S.D.		MAIN	COMPSTAT	POSITION	INTERACTION
7. PREDICTED QUALITY (PRLLQL)	49	Low	-1.34	2.71	F	3.67	4.69	.789	2.84
	24	Mod	-0.72	2.27	df	3,91	2,91	1,91	2,91
	24	High	0.63	2.35	p <	.015	.01	.377	.064
	97	Total	-0.70	2.62					

VARIABLE		
	Low	F
	Mod	df
	High	p
	Total	

VARIABLE		
	Low	F
	Mod	df
	High	p
	Total	

TABLE 45

DISCRIMINANT FUNCTION CLASSIFICATION ANALYSES, N=97, PRETEST

GROUP MEMBERSHIP

	ACTUAL	N	PREDICTED GROUP 1.	2.	3.
ANALYSIS #1	1. LOW	49	37 75.5%	10 20.4%	2 4.1%
	2. MOD	24	4 16.7%	18 75.0%	2 8.3%
	3. HIGH	24	0	5 20.8%	19 79.2%
			% of "Grouped" cases correctly classified = 76.29%		
ANALYSIS #2			1.	2.	3.
	1. LOW	49	32 65.3%	11 22.4%	6 12.2%
	2. MOD	24	6 25.0%	10 41.7%	8 33.3%
	3. HIGH	24	4 16.7%	2 8.3%	18 75.0%
			% of "Grouped" cases classified correctly = 61.86%		
ANALYSIS #3			1.	2.	3.
	1. LOW	49	34 69.4%	8 16.3%	7 14.3%
	2. MOD	24	4 16.7%	12 50.0%	8 33.3%
	3. HIGH	24	2 8.3%	6 25.0%	16 66.7%
			% of "Grouped" cases classified correctly = 63.92%		
ANALYSIS #4			1.	2.	3.
	1. LOW	49	39 79.6%	10 20.4%	0
	2. MOD	24	2 8.3%	20 83.3%	2 8.3%
	3. HIGH	24	0	4 16.7%	20 83.3%
			% of "Grouped" cases correctly classified = 81.4%		

TABLE 47

Inter-Domain Correlations PRETEST (N=97)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	—	.17	.17	.11	.12	.35	.21	.22	.03	.04	.01	-.10	.10
2		—	-.03	.07	.02	.19	.20	.17	.18	.06	.11	-.11	.16
3			—	.01	.20	.23	.002	-.01	.20	.12	.23	.05	.24
4				—	.44	.31	.33	.08	.04	.15	.23	.08	.17
5					—	.47	.35	-.35	.19	.04	.33	.06	.29
6						—	.62	.38	.06	.18	.31	.02	.34
7							—	.60	.15	.19	.30	.34	.38
8								—	.13	.19	.15	.21	.23
9									—	-.20	.65	.30	.51
10										—	.33	.66	.43
11											—	.77	.83
12												—	.79
13													—

(N.B. When r > .20, p < .05,
and when r > .26, p < .01.)

ABLES: 1) PERCEPTION: SPEED; 2) PERCEPTION: COMPOSITE ACCURACY; 3) COGNITION: COGSUM; 4) STL,
positional Usage; 5) STL, Efficiency; 6) STL, Organisation; 7) STL, Quality; 8) STL, Quantity;
LL, Efficiency; 10) LL, Quantity; 11) LL, Organisation; 12) LL, Quality; 13) LL, Discrimination

TABLE 48: Inter-Domain Correlations, Post-test (N=49).

(N.B. When $r > .28$, $p < .05$, and $r > .36$, $p < .01$.)

```
                              SAMPLE CORRELATION MATRIX
                    ------------------------------------------------------------

                1          2          3          4          5          6          7          8          9         10
              MIPS       PACC       COGS       SPUP       STLQ       FMSF       SPUP        X2        PRST       LLLW

 1  PTPS    1.000000
 2  PACC    0.233629   1.000000
 3  COGS   -0.185242   0.335358   1.000000
 4  SPUP   -0.025480   0.214011   0.203112   1.000000
 5  STLQ    0.198567   0.403149   0.065347   0.281864   1.000000
 6  FMSF    0.087260   0.074145  -0.034449   0.249431   0.180202   1.000000
 7  SPUP    0.145062   0.145111   0.029474   0.346910   0.260985   0.551753   1.000000
 8  X2     -0.022054  -0.161981  -0.040691   0.171132  -0.486894   0.394479   0.704009   1.000000
 9  PRST    0.061942   0.276970   0.061503   0.410575   0.368757   0.468464   0.857950   0.522772   1.000000
10  LLLQ   -0.100779   0.054036   0.267995   0.347128   0.032625  -0.048963   0.133508   0.115497   0.119607   1.000000
11  LLLT    0.085519   0.038417   0.188453  -0.185814  -0.171284  -0.127852   0.174365   0.244332   0.125160  -0.197244
12  ALLU   -0.055036  -0.012972   0.386395   0.265677  -0.084113   0.021123   0.248700   0.284184   0.145193   0.711312
13   POS    0.103830   0.022114   0.270339   0.095649   0.056660  -0.026231   0.257671   0.183207   0.205299   0.543820
14  PRLL    0.083154   0.007557   0.233292   0.206567   0.152418   0.031188   0.224951   0.100818   0.185986   0.752398
15  SVCD    0.149895   0.136142   0.326913   0.348235  -0.045481   0.120295   0.317591   0.312476   0.197407   0.663376


               11         12         13         14         15
              LLLT       ALLU        POS       PRLL       SVCD

11  LLLT    1.000000
12  ALLU    0.141427   1.000000
13   POS    0.506244   0.662306   1.000000
14  PRLL    0.215443   0.728857   0.917411   1.000000
15  SVCD    0.255594   0.746409   0.534240   0.518421   1.000000
```

VARIABLES:

1. Perception: Speed
2. " : Composite Accuracy
3. Cognition: COGSUM
4. STL: Prepositional Accuracy
5. " : Efficiency
6. " : Organisation
7. " : Quality
8. " : Quantity
9. " : Predicted Quality
10. LL: Efficiency
11. " : Quantity
12. " : Organisation
13. " : Quality
14. " : Predicted Quality
15. " : Discrimination

TABLE 49

Factor Analyses of Language Domain Variables, Pretest (N=97)

VARIMAX ROTATED FACTOR MATRIX

VARIABLES		FACTORS			
		1	2	3	4
STL	1. Prepositional Accuracy	.036	.755	.013	.091
	2. Efficiency	.217	.783	-.458	-.053
	3. Organization	.139	.739	.429	.078
	4. Quantity	.115	.018	.956	.081
LL	5. Discrimination	.826	.173	.119	.387
	6. Organization	.898	.214	.025	.242
	7. Efficiency	.870	-.026	.053	-.387
	8. Quantity	.118	.074	.092	.957

OBLIQUE ROTATION FACTOR PATTERN

VARIABLES	FACTORS			
	1	2	3	4
1	-.042	.759	.055	.005
2	.167	.774	-.067	-.467
3	.050	.742	.018	.422
4	.074	.019	.027	.954
5	.790	.077	.372	.069
6	.873	.122	.227	-.019
7	.911	-.074	-.401	.048
8	.043	-.003	.962	.031

TABLE 50

Factor Analyses of Language Domain Variables, Post-test (N=97)

CORRELATION MATRIX

VARIABLES		1	2	3	4	5	6	7	8
STL	1. Prepositional Accuracy	—	.29	.26	.11	.17	.16	.20	-.13
	2. Efficiency		--	.15	-.51	-.006	-.03	.05	-.12
	3. Organization			--	.35	.09	.02	-.01	.02
	4. Quantity				—	.26	.23	.04	.26
LL	5. Discrimination					--	.80	.41	.28
	6. Organization						--	.62	.11
	7. Efficiency							--	.34
	8. Quantity								—

VARIMAX ROTATED FACTOR MATRIX

VARIABLES	1	2	3	4
1.	.195	.214	.698	-.194
2.	.024	.890	.306	.001
3.	-.054	-.137	.844	.094
4.	.183	-.815	.388	.173
5.	.891	-.047	.106	.252
6.	.947	-.073	.035	-.004
7.	.700	-.006	.028	-.588
8.	.147	-.109	-.042	.927

TABLE M1

Multivariate Analyses of Training Effects for the Four Variables Representing Each Domain, for EXPGROUPS (N = 49)

SCORES: GAIN SCORES

VARIABLES:	I $\overline{X}$	I S.D.	II $\overline{X}$	II S.D.	III $\overline{X}$	III S.D.	IV $\overline{X}$	IV S.D.
1. PERCEPTION Total # Choices	0.54	3.38	-0.58	2.27	-1.09	2.88	-2.15	3.53
2. COGNITION 3-score	0.15	1.46	1.83	1.95	1.73	1.68	0.46	1.76
3. STL COMPOSITE COMPETENCE	20.45	10.42	19.19	10.45	5.91	15.05	15.24	15.62
4. LL Picture Boxes Discrimination	2.38	2.87	3.25	2.30	5.27	3.13	2.54	2.26

CONTRASTS SIMPLE	VARS. "BIG 4"	Mean Square	MANOVA (df= 4,42) F	p<	UNIVARIATE (df= 1,45) F	p<	STEP-DOWN F	p<
D. I vs. IV			1.96	.12				
	1	32.48			3.42	.07	3.42	.07
	2	12.67			4.29	.04	2.24	.14
	3	433.96			2.54	.12	1.07	.31
	4	14.37			2.04	.16	1.01	.32
E. II vs. IV			1.73	.16				
	1	9.39			0.99	.33	0.99	.33
	2	5.01			1.70	.19	2.77	.10
	3	541.76			3.17	.08	2.81	.10
	4	2.35			0.33	.57	0.28	.60
F. III vs. IV			3.08	.026				
	1	6.73			0.71	.40	0.71	.40
	2	9.55			3.24	.078	4.54	.038
	3	518.97			3.03	.088	2.59	.11
	4	44.55			6.34	.015	3.71	.06

TABLE M2

Multivariate Analyses of the Four Variables from Each Domain, for EXPGROUPS (N = 49)

SCORES: GAIN SCORES

VARIABLES: See Table M1

CONTRASTS HELMERT	VARS. "BIG 4"	Mean Square	MANOVA (df= 4,42)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p <	F	p <	F	p <
			1.95	.12				
A. Exptl. vs. Control Groups	1	31.65			3.34	.07	3.34	.07
	2	5.13			1.74	.19	3.61	.06
	3	1.14			.007	.94	.006	.94
	4	9.88			1.40	.24	.837	.37
			2.74	.04				
B. I vs. II + III	1	15.46			1.63	.21	1.63	.21
	2	22.03			7.47	.009	5.69	.02
	3	480.73			2.81	.10	1.25	.27
	4	27.90			3.96	.05	1.97	.17
			2.08	.10				
C. II vs. III	1	1.48			0.16	.69	0.16	.69
	2	0.06			0.02	.88	0.08	.78
	3	1012.82			5.92	.019	5.66	.02
	4	23.48			3.34	.07	2.25	.14

TABLE M3

Multivariate Analyses of Training Effects

for Perception and Cognition, for EXPGROUPS (N = 49)

SCORES: Unadjusted Post-test Scores (see Table 2-38)

VARIABLES: 1. Perception: SPEED
2. Perception: COMPOSITE ACCURACY
3. Cognition: COGSUM

CONTRASTS HELMERT	VARS. Perception x Cognition	Mean Square	MANOVA (df= 3,43)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Groups			0.59	.62				
	1	9.12			0.48	.49	0.48	.49
	2	2.41			0.05	.82	0.16	.69
	3	16.23			1.63	.21	1.01	.32
B. I vs. II + III			2.36	.08				
	1	0.44			0.02	.88	.02	.88
	2	50.32			1.04	.31	.99	.32
	3	69.03			6.93	.01	6.25	.016
C. II vs. III			0.76	.52				
	1	4.15			0.22	.64	0.22	.64
	2	22.39			0.46	.50	0.33	.57
	3	8.12			0.82	.37	1.45	.24

TABLE M4

Multivariate Analyses of Training Effects for Perception and Cognition with the LL Discriminability Variable, for EXPGROUPS (N = 49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

VARIABLES:
1. Perception: SPEED
2. Perception: COMPOSITE ACCURACY
3. LL: DISCRIMINABILITY INDEX
4. Cognition: COGSUM

CONTRASTS HELMERT	VARS. Perception x Cognition x LL DISCR.	Mean Square	MANOVA (df=4,42)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Groups	1	9.12	3.06	.027	0.48	.49	0.48	.49
	2	2.41			0.05	.82	0.16	.69
	3	62.27			11.58	.001	11.71	.001
	4	16.23			1.63	.21	0.02	.88
B. I vs. II + III	1	0.44	1.79	.15	0.02	.88	0.02	.88
	2	50.32			1.04	.31	0.99	.32
	3	1.23			0.23	.63	0.12	.73
	4	69.03			6.93	.01	5.93	.019
C. II vs. III	1	4.15	0.47	.76	0.22	.64	0.22	.64
	2	22.39			0.46	.50	0.33	.57
	3	0.00			0.00	1.0	0.02	.90
	4	8.12			0.82	.37	1.32	.26

TABLE M5

Multivariate Analyses of Training Effects

for 5 STL Variables, for EXPGROUPS (N = 49)

SCORES: GAIN SCORES

VARIABLES:

	EXPGROUPS I		II		III		IV	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
1. Quality	5.54	6.29	5.08	3.63	2.91	5.01	3.77	2.83
2. Inefficiency	-1.46	2.88	-0.33	7.15	-0.09	6.83	0.15	4.51
3. Efficiency	17.02	12.56	14.95	12.70	5.41	17.93	9.62	14.56
4. Organisation	4.46	3.69	4.67	3.47	1.36	4.82	3.46	4.22
5. Composite Competence	20.45	10.42	19.19	10.45	5.91	15.05	15.24	15.62

CONTRASTS ARBITRARY	VARS. STL	Mean Square	MANOVA (df= 5,41)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
1. I + II vs. IV			2.38	.05				
	1	55.18			2.57	.11	2.57	.11
	2	11.76			0.39	.54	0.56	.46
	3	951.70			4.54	.039	4.21	.046
	4	59.51			3.61	.06	1.64	.21
	5	1107.94			6.48	.01	2.32	.14

TABLE M6

Multivariate Analysis of Training Effects for 4 LL Variables, for EXPGROUPS (N=49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

VARIABLES:

1. Organisation
2. Efficiency
3. Quantity
4. Discrimination

CONTRASTS HELMERT	VARS. LL	Mean Square	MANOVA (df= 4,42)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Grps			3.34	.018				
	1	19.00			5.99	.018	5.99	.018
	2	0.12			0.17	.69	3.11	.085
	3	219.19			3.67	.06	0.57	.45
	4	62.27			11.58	.001	3.01	.09
B. I vs. II + III			0.60	.66				
	1	0.26			0.08	.78	.08	.78
	2	0.003			0.004	.95	0.04	.85
	3	79.66			1.33	.25	1.94	.17
	4	1.23			0.23	.63	0.38	.54
C. II vs. III			0.46	.76				
	1	1.78			0.56	.46	0.56	.46
	2	0.09			0.12	.73	0.07	.79
	3	20.75			0.35	.56	0.77	.39
	4	0.00			0.00	1.00	0.49	.49

TABLE M7

Multivariate Analyses of Training Effects for Cognitive and STL Variables, for EXPGROUPS (N=49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

VARIABLES:
1. Cognition: COGSUM
2. STL: Organisation
3. STL: Efficiency
4. STL: Prepositional Usage

CONTRASTS HELMERT	VARS. COGNITION x STL	Mean Square	MANOVA (df= 4,42)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p <	F	p <	F	p <
A. Exptl. vs. Control Grps.			0.68	.61				
	1	16.23			1.63	.21	1.63	.21
	2	0.08			0.009	.92	0.02	.89
	3	12.49			0.98	.33	1.14	.29
	4	0.009			0.01	.92	0.02	.89
B. I vs. II + III			2.81	.037				
	1	69.03			6.93	.01	6.93	.01
	2	29.97			3.27	.077	2.50	.12
	3	7.69			0.60	.44	0.28	.60
	4	1.38			1.63	.21	1.37	.25
C. II vs. III			0.99	.43				
	1	8.12			0.82	.37	.82	.37
	2	23.31			2.54	.12	2.35	.13
	3	0.05			.004	.95	.096	.76
	4	0.43			0.50	.48	0.75	.39

5

TABLE M8

Multivariate Analyses of Training Effects for Cognitive and LL Variables, for EXPGROUP (N=49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

VARIABLES:

1. Cognition: COGSUM
2. LL: Organization
3. LL: Efficiency
4. LL: Quantity
5. LL: Discrimination

CONTRASTS HELMERT	VARS. Cognition x LL	Mean Square	MANOVA (df= 5,41)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Grps.			2.61	.039				
	1	16.23			1.63	.21	1.63	.21
	2	19.00			5.99	.018	4.24	.04
	3	0.12			0.17	.69	3.02	.089
	4	219.19			3.67	.06	0.51	.48
	5	62.27			11.58	.001	2.92	.09
B. I vs. II + III			2.08	.087				
	1	69.03			6.93	.01	6.93	.01
	2	0.26			0.08	.78	0.53	.47
	3	.003			.004	.95	0.02	.88
	4	79.66			1.33	.25	2.83	.10
	5	1.23			0.23	.63	0.22	.64
C. II vs. III			0.60	.70				
	1	8.12			0.82	.37	0.82	.37
	2	1.78			0.56	.46	1.36	.25
	3	0.09			0.12	.73	0.07	.79
	4	20.75			0.35	.56	0.45	.51
	5	0.00			0.00	.00	0.41	.52

TABLE M9

Multivariate Analyses of Training Effects for Efficiency Measures in STL and LL Domains, for EXPGROUPS (N=49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

VARIABLES:
1. LL: Organisation
2. LL: Efficiency
3. LL: Quantity
4. LL: Discrimination
5. STL: Prepositional Usage
6. STL: Efficiency
7. STL: Organisation

CONTRASTS HELMERT	VARS. STL x LL	Mean Square	MANOVA (df= 7,39)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p <	F	p<	F	p <
A. Exptl. vs. Control Grps.			1.88	.099				
	1	19.00			5.99	.018	5.99	.018
	2	0.12			0.17	.69	3.11	.08
	3	219.19			3.67	.06	0.57	.45
	4	62.27			11.58	.001	3.01	.09
	5	.009			0.01	.92	0.55	.46
	6	12.49			0.98	.33	.002	.96
	7	.084			.009	.92	.003	.96
B. I vs. II + III			1.25	.30				
	1	0.26			0.08	.78	0.08	.78
	2	.003			.004	.95	0.04	.85
	3	79.66			1.33	.25	1.94	.17
	4	1.23			0.23	.63	0.38	.54
	5	1.38			1.63	.21	3.23	.08
	6	7.69			0.60	.44	0.03	.85
	7	29.97			3.27	.077	2.68	.11
C. II vs. III			0.75	.63				
	1	1.78			0.56	.46	0.56	.46
	2	0.09			0.12	.73	0.07	.79
	3	20.75			0.35	.56	0.77	.39
	4	0.00			0.00	1.00	0.49	.49
	5	0.43			0.50	.48	0.66	.42
	6	0.05			.004	.95	.003	.95
	7	23.31			2.54	.12	2.39	0.13

TABLE M10

Multivariate Analyses of Training Effects for Quantity and Quality in the STL and LL Domains, for EXPGROUPS (N=49)

SCORES: Unadjusted Post-Test Scores (see Tables 2-38)

VARIABLES:			
STL	1. Quantity	**LL**	5. Quantity
	2. Quality		6. Quality
	3. Organisation		7. Organisation
	4. Prepositional Usage		8. Discrimination

CONTRASTS HELMERT	VARS. STL x LL	Mean Square	MANOVA (df= 8,38)		UNIVARIATE (df=1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Grps.			1.50	.19				
	1	114.02			2.31	.14	2.31	.14
	2	20.52			1.15	.29	0.0	1.00
	3	0.08			.009	.92	0.34	.57
	4	,009			0.01	.92	0.10	.75
	5	219.19			3.67	.06	1.78	.19
	6	63.63			4.42	.04	1.15	.29
	7	19.00			5.99	.018	2.23	.14
	8	62.27			11.58	.001	3.47	.07
B. I vs. II + III			1.21	.32				
	1	20,61			0.42	.52	0.42	.52
	2	38.77			2.17	.15	1.98	.17
	3	29.97			3.27	.077	1.31	.26
	4	1.38			1.63	.21	0.39	.54
	5	79.66			1.33	.25	2.20	.15
	6	0.02			.001	.97	1.27	.27
	7	0.26			0.08	.78	0.05	.83
	8	1.23			0.23	.63	1.95	.17
C. II vs. III			0.73	.66				
	1	10.67			0.22	.64	0.22	.64
	2	1.27			0.07	.79	.007	.93
	3	23.31			2.54	.12	2.85	.099
	4	0.43			0.50	.48	0.78	.38
	5	20.75			0.35	.56	0.21	.65
	6	1.03			0.07	.79	0.12	.73
	7	1.78			0.56	.46	1.48	.23
	8	0.00			0.00	1.00	0.38	.54

TABLE M11

Multivariate Analyses of Training Effects for Predicted Quality Measures in the STL and LL Domains for EXPGROUPS (N=49)

SCORES: Unadjusted Posttest Scores (see Tables 2-38)

VARIABLES:

STL	1. Prepositional Usage	LL	4. Organisation
	2. Organisation		5. Predicted Quality
	3. Predicted Quality		6. Discrimination

CONTRASTS ELEMENT	VARS. STL x LL	Mean Square	MANOVA (df= 6.40)		UNIVARIATE (df= 1,45)		STEP-DOWN	
			F	p<	F	p<	F	p<
A. Exptl. vs. Control Grps.			2.01	.087				
	1	.009			0.01	.92	.01	.92
	2	0.08			.009	.92	.02	.90
	3	.009			.001	.98	.001	.97
	4	19.00			5.99	.018	6.23	.017
	5	25.16			2.54	.12	.05	.83
	6	62.27			11.58	.001	5.23	.03
B. I vs. II +III			0.81	.57				
	1	1.38			1.63	.21	1.63	.21
	2	29.97			3.27	.077	2.23	.14
	3	13.50			0.78	.38	0.06	.80
	4	0.26			0.08	.78	0.27	.61
	5	2.66			0.27	.61	0.20	.66
	6	1.23			.23	.63	0.72	.40
C. II vs. III			0.97	.46				
	1	0.43			0.50	.48	.50	.48
	2	23.31			2.54	.12	3.23	.079
	3	0.56			0.03	.86	.06	.80
	4	1.78			0.56	.46	1.01	.32
	5	0.01			0.001	.97	.56	.46
	6	0.00			0.0	1.00	.63	.43

TABLE M12

Multivariate Analyses of Training Effects, Grand MANOVA for all Domains, for EXPGROUPS (N=49)

SCORES: Unadjusted Post-test Scores (see Tables 2-38)

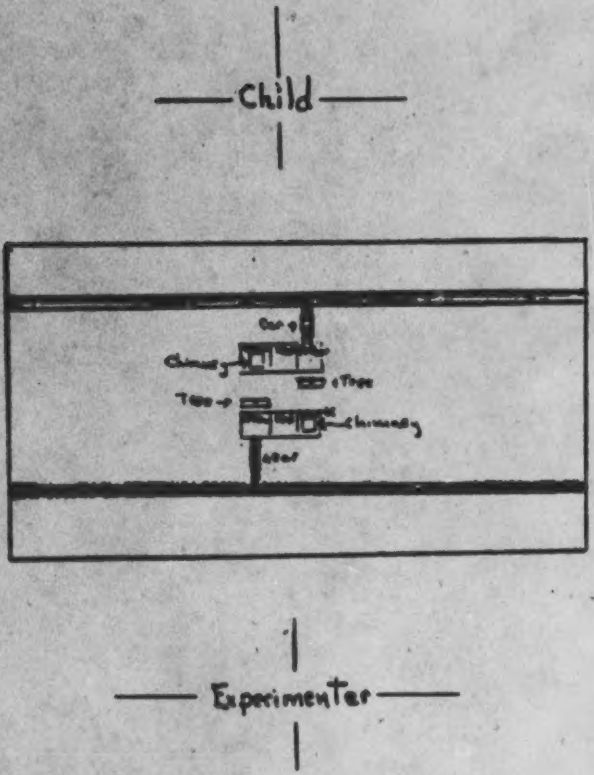
VARIABLES:

1. LL Organisation
2. STL Organisation
3. STL Prepositional Accuracy
4. LL Quantity
5. STL Quantity
6. LL Efficiency
7. STL Efficiency
8. LL Discrimination
9. Perception/Accuracy
10. Perception/Speed
11. Cognition/Cogsum

CONTRASTS HELMERT	VARS. Perceptual, Cognitive, STL & LL	Mean Square	MANOVA (df=11,35) F	MANOVA (df=11,35) p<	UNIVARIATE (df=1,45) F	UNIVARIATE (df=1,45) p<	STEP-DOWN F	STEP-DOWN p<
A. Exptl. vs. Control Grps.			1.24	.297				
	1	19.00			5.99	.02	5.99	.018
	2	0.08			.009	.92	.002	.97
	3	.009			0.01	.92	.56	.46
	4	219.19			3.67	.06	1.75	.19
	5	114.02			2.31	.14	0.19	.66
	6	0.12			0.17	.69	1.19	.28
	7	12.49			0.98	.33	.002	.96
	8	62.27			11.58	.001	3.20	.08
	9	2.41			0.05	.82	0.01	.91
	10	9.12			0.48	.49	1.32	.26
	11	16.23			1.63	.21	.002	.96
B. I vs. II + III			1.61	.14				
	1	0.26			0.08	.78	0.08	.78
	2	29.97			3.27	.077	3.21	.08
	3	1.38			1.63	.21	0.85	.36
	4	79.66			1.33	.25	2.47	.12
	5	20.61			0.42	.52	0.21	.65
	6	.003			.004	.95	0.58	.45
	7	7.69			0.60	.44	0.03	.85
	8	1.23			0.23	.63	1.49	.23
	9	50.32			1.04	.31	1.67	.20
	10	0.44			0.02	.88	0.02	.89
	11	69.03			6.93	.01	5.98	.019
C. II vs. III			0.68	.75				
	1	1.78			0.56	.46	0.56	.46
	2	23.31			2.54	.12	2.41	.13
	3	0.43			0.50	.48	1.72	.20
	4	20.75			0.35	.56	0.55	.46
	5	10.67			0.22	.64	0.03	.87
	6	0.09			0.19	.73	.009	.92
	7	0.05			.004	.95	.003	.95
	8	0.00			0.00	1.00	0.30	.59
	9	22.39			0.46	.50	0.93	.34
	10	4.15			0.22	.64	0.15	.70
	11	8.12			0.82	.37	1.36	.25

Figure 1.

Cognitive Task #1 - Block Construction


Child
Car
Chimney
Tree
Tree
Chimney
Car
Experimenter

Training

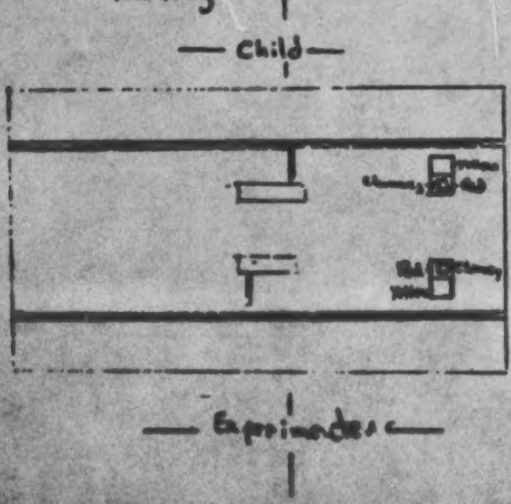

Child
Yellow
Chimney
Red
Red
Chimney
Yellow
Experimenter

Figure 2.

Cognitive Task #2 - House Pictures.

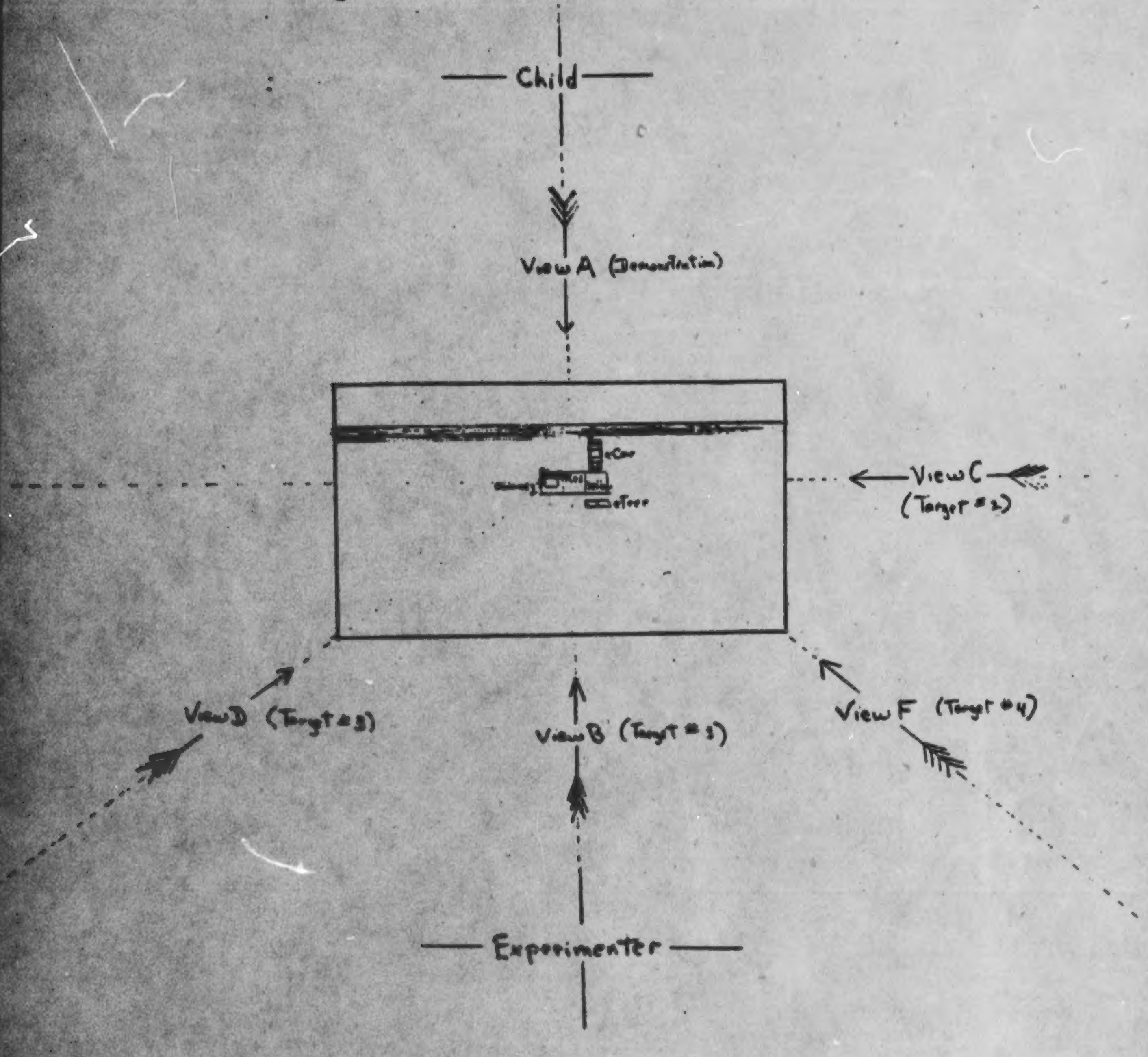
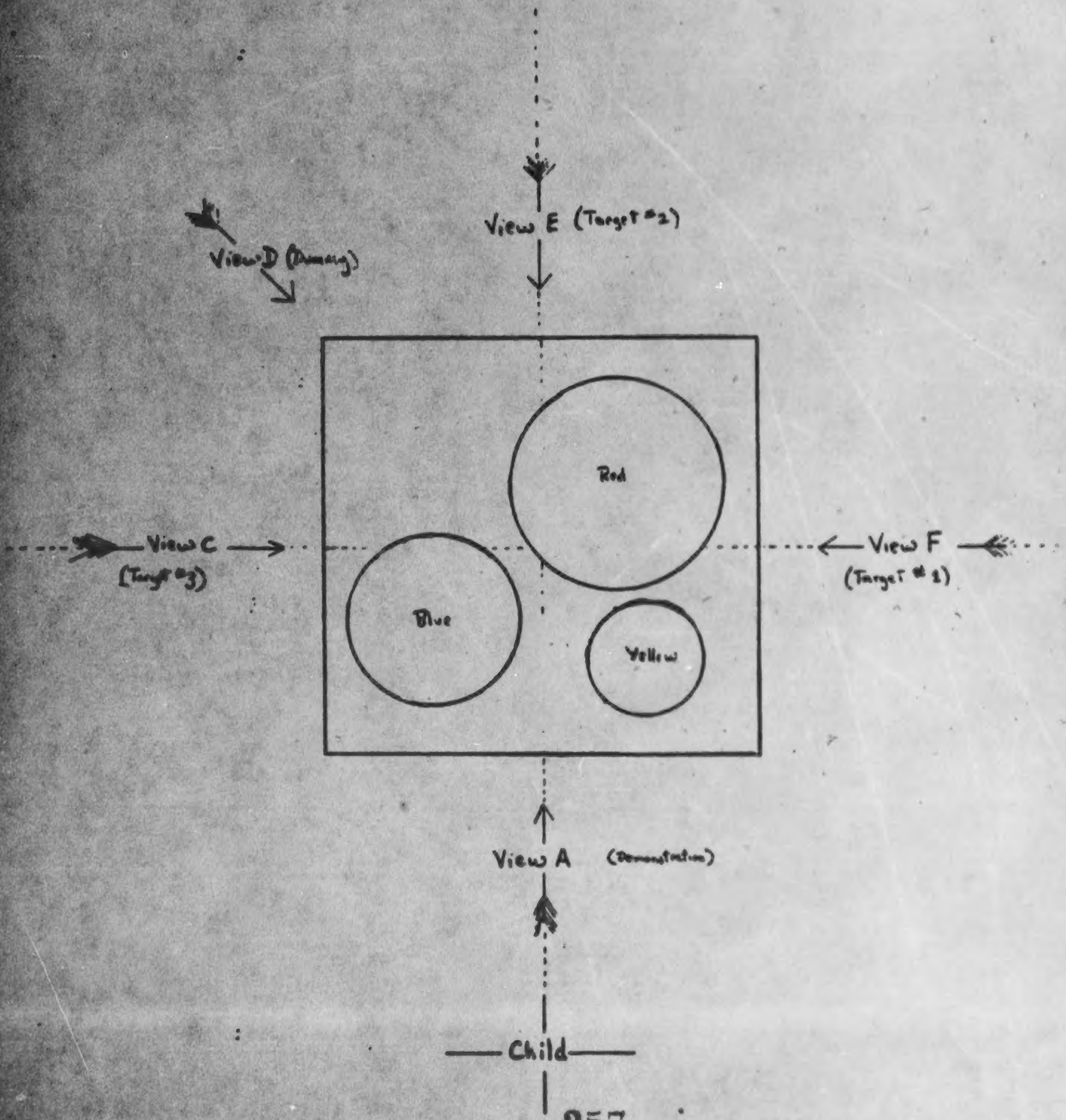

Child
View A (Demonstration)
Car
Tree
View C
(Target #2)
View D (Target #3)
View B (Target #1)
View F (Target #4)
Experimenter

Figure 3.

Cognitive Task #3 - Mountains


View E (Target #2)
View D (Dummy)
Red
View C
(Target #3)
View F
(Target #1)
Blue
Yellow
View A (Demonstration)
Child

APPENDIX A: Perceptual Tasks Score Sheet

Spontaneous Search

Name: ____________	Class: ____________
Date: ____________	Subject #: ____________

Demonstration:

a) Bottle (4)	1. ________	2. ________	
	3. ________	4. ________	

Practice:

A. Apple (3)	1. ________	2. ________	
	3. ________	4. ________	
B. Flower (3)	1. ________	2. ________	
	3. ________	4. ________	

Test:

1. Scissors (6)	1. ________	2. ________	3. ________
	4. ________	5. ________	6. ________
2. Tree (2)	1. ________	2. ________	3. ________
	4. ________	5. ________	6. ________
3. Lamp (5)	1. ________	2. ________	3. ________
	4. ________	5. ________	6. ________
4. Bear (1)	1. ________	2. ________	3. ________
	4. ________	5. ________	6. ________
5. Cowboy (4)	1. ________	2. ________	3. ________
	4. ________	5. ________	6. ________

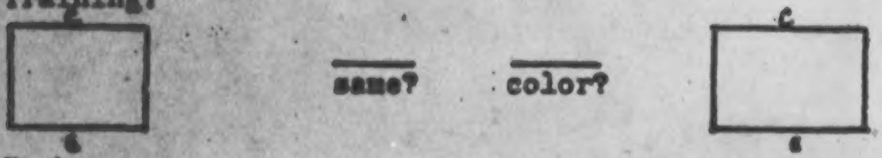
APPENDIX B: Cognitive Tasks Score Sheet

NAME______________________ DATE__________________

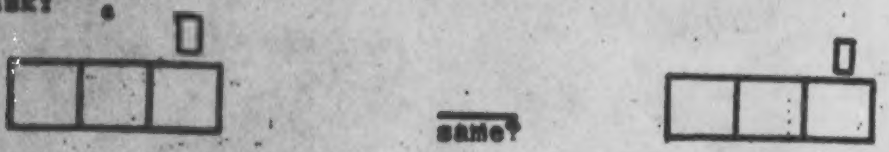
CLASS_____________________ SUBJECT #______________

ii-1 Matching a Structure

Training:

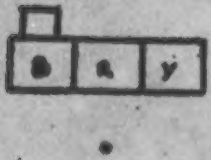

c
s
same?
color?
c
s

Task:

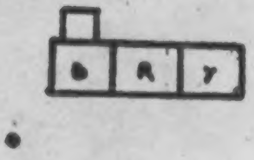


ii-2 Choosing a Position

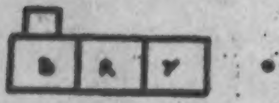
View 1 (E):



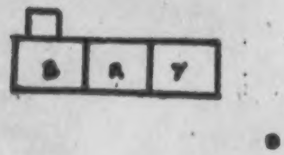
View 3 (D):



View 2 (C):



View 4 (F):



ii-3 Choosing a Photo

1. View F photo______
2. View E photo______
3. View C photo______

APPENDIX C: Language Tasks: STL Domain

Comprehension of Prepositions

Subject ______________________ Class ______________________

Date ______________________ Order of presentation # ______

1. Put the block on the big desk.

2. Put the block in back of the little desk.

3. Put the block between the desks.

4. Put the block in the little desk.

5. Put the block under the big desk.

6. Put the block in front of the little desk.

APPENDIX D: Language Tasks: Spatio-Temporal Language Test Items; Experimenter's Script for demonstrating action sequedces.

PRACTICE:

A: Put the seal on the car.

B: Put the soldier in front of the house.

TASKS:

Encoder #1A: Drive the car behind the house and put the seal in the house.

#2A: Put the red block in front of the yellow block and put the soldier on both blocks.

#1B: Put the hippo on the car, and drive under the bridge and up to the red block, and put the hippo under the soldier.

#2B: Take the elephant and hop him until he is between the trees and put the hippo behind the elephant.

#1C: Put the elephant in front of one tree and the hippo in front of the other tree, and drive the car between them.

#2C: Put the elephant and hippo in the car and drive it to the house. Put the elephant in the house and put the hippo under the soldier.

Order of Prepositions:	Encoder #1	Encoder #2
	behind	in front of
	in	on
	on	between
	under	behind
	in front of	in
	between	under

APPENDIX E: Language Tasks: STL ANSWER SHEET

KN #1 ________________ CLASS ________________

KN #2 ________________ CLASS ________________

DATE ________________

Practice:

A: Seal ________________

B: Soldier ________________

Task:

KN #1-a: Car ________________

Seal ________________

KN #2-a: Red block ________________

Soldier ________________

KN #1-b: Car ________________

Hippo ________________

KN #2-b: Elephant ________________

Hippo ________________

KN #1-c: Elephant ________________

Hippo ________________

Car ________________

KN #2-c: Elephant ________________

Hippo ________________

APPENDIX F: **Language Tasks: Logical Language Test Items.**

Answer sheet

Subject A ____________________

" B ____________________

Date ____________________

Class ____________________

Dot indicates position of target in Decoder's array.

PRODUCTION BOARD

I. (soldier)	Enc. A	☐ ☐ ☐ ☐ ■	Enc. B	☐ ☐ ☐ ■ ☐
II. (shades)	Enc. B	■ ☐ ☐ ☐ ☐	Enc. A	☐ ■ ☐ ☐ ☐

PICTURE BOXES

I. (dog)	Enc. A	☐ ■ ☐ ☐ ☐	Enc. B	☐ ■ ☐ ☐ ☐
II. (games)	Enc. B	☐ ☐ ☐ ☐ ■	Enc. A	☐ ☐ ☐ ☐ ■
III. (babies)	Enc. A	☐ ☐ ☐ ■ ☐	Enc. B	☐ ☐ ☐ ☐ ■

CARTOON BOXES

I. (Mutt&J.)	Enc. B	☐ ☐ ■ ☐	Enc. A	☐ ■ ☐ ☐
II. (snow)	Enc. A	☐ ☐ ■ ☐ ☐	Enc. B	☐ ☐ ☐ ■ ☐
III. (prince)	Enc. B	☐ ■ ☐ ☐ ☐	Enc. A	☐ ☐ ☐ ■ ☐

APPENDIX G: Language Tasks: Excerpts from Test Transcripts, Pre-Training and Post-Training, showing responses to STL and LL Tasks, with examples of the scoring system.

Task: STL/Action-Sequence
Subject: #91
Status: Low, Experimental

Pretest	Posttest
TEST DEMONSTRATION A (adequate description)	TEST DEMONSTRATION A
Put the red block in front of the yellow block and put the soldier on both blocks.	Drive the car behind the house and put the seal in the house.
Subject's response to A:	Subject's response to A:
RC: The yellow brick . . . near . . . the wall . . . she put the red brick in front of it and she put the man on top of it.	RC: You know the house where the soldier is? WN: Yeah. RC: Well, take the orange car and drive it down the road, and make it go to the left, and then make it go to the right, and then make it go into that road in back of the house, and then make it stop, and then put the seal inside.

* * * * *

Pretest	Posttest
TEST DEMONSTRATION B	TEST DEMONSTRATION B
Take the elephant and hop him until he is between the trees and put the hippo behind the elephant.	Put the hippo on the car, and drive under the bridge and up to the red block, and put the hippo under the soldier.
Subject's response to B:	Subject's response to B:
RC: The two trees right? You know that little road coming in . . . in between them?	RC: You know the orange car? WN: Yeah.

Pretest

MO: Yeah

RC: She put the elephant . . . right . . . you know that road that come down?

MO: What road?

RC: This road.

E: No, tell her.

RC: You know the road that lead up to the . . . house?

MO: Yeah.

RC: You know that road that you have to go down to get to the house?

MO: Yeah.

RC: Well, that road . . . she put the elephant . . ..right where the go into the bridge . . . and then the hippopotamus in back.

Posttest

RC: Well, take the hippo and put it on the orange car, then make it come out the road and go on to that really long one, then make it go under through the bridge, and than make it stop in front, over by, about 1/2 inch away from the . . . ah . . . the red block and then put the hippopotamus underneath the soldier.

* * * * *

TEST DEMONSTRATION C

Put the elephant and hippo in the car and drive it to the house. Put the elephant in the house and put the hippo under the soldier.

Subject's Response to C:

RC: You know the one . . . where the . . . seal is in the house?

MO: Yeah.

RC: Well, put the elephant in there . . . you know the one where . . . the hippopotamus is under . . . under the soldier?

MO: Yeah.

RC: Do that . . . and put the car in front of the house.

TEST DEMONSTRATION C

Put the elephant in front of one tree and the hippo in front of the other tree, and drive the car between them.

Subject's response to C:

RC: You know the orange car?

WN: Yeah.

RC: Drive it all the way out till you get to the . . . don't turn ujst keep on going straight till you get them out? No, make them go frontwards, and make them stop at the end of the blue line. Then put the elephant near the . . . ah . . . right side, the right one near the tree. Put the elephant near the right side of the tree right in front of it and put the hippopotamus in front of the left one.

SCORING

Pretest

Propositions rated + = 10

Propositions rated 0 = 11

Propositions rated - = 2

Total FM score = 9

Posttest

Propositions rated + = 28

Propositions rated 0 = 14

Propositions rated - = 0

Total FM score = 17

Task: LL/Picture Description
Subject: #91
Status: Low, Experimental

Pretest

PRODUCTION BOARD I

RC: There's a lady with black and, uh, she has a green shirt on and a white necklace and a green jacket and a man has a white jacket on with a different colored tie.

PRODUCTION BOARD II

RC: The sick looking boy. He's a man. The, uh, nice big tie.

AM: What?

RC: There's a man that, um, mustache, and there's a, um, and he has a lady with kind of brown hair in a, um, a bow up near her neck.

Posttest

PRODUCTION BOARD I

RC: The woman has a white dress on, she's black...and the boy...has a sport jacket on...and a black tie... and he has two red stripes on the side...and a white kind of like collar...and he has, um...he has a special kind of hair-do...and the girl has glasses and a necklace.

* * * * *

PRODUCTION BOARD II

RC: There's a white man and a...a black woman. There's a black man and a white lady...and they are right...they're right across from the Chinese people.

E. Hers are in different places.

RC: And, and the girl has a pink ...kind of like bodyshirt, and a red dress...and the man has like a kind of swervy tie...and, and, uh, like things coming down from his hair like, and he cut part of, all of his beard, and he has a mustache...and he has a sport jacket on, and he has black pants.

* * * * *

Pretest	Posttest
PICTURE BOXES I	PICTURE BOXES I
RC: She has, um, a bracelet and a watch and she has earrings. She has lipstick. She has like dark eyebrows.	RC: They're facing nose to nose and the girl is bending down...and she has a watch on and she has...she has a ring on her middle finger, and she has, you know like, where you have this, um, bracelet and it says your name on it? BF: Yeah. RC: She has one of those on...and she has a watch right underneath her. And you can see her feet...and she has a bracelet on both hands.
* * * * *	
PICTURE BOXES II	PICTURE BOXES II
RC: They're play- the boy's laying down. The girl's sitting up and it's the boy's turn and uh, the um, the girl has flowers on her shirt, and she has blue pants, and she has like kind of white at the end of her fingernails.	RC: This is easy. The two kids are playing Gnip-Gnop...and there's, and there's a tugboat...on the side that the boy's laying down.
* * * * *	
PICTURE BOXES III	PICTURE BOXES III
RC: There's a like, uh, a little boy with black, um, like a black thing on, like pants and shirt. AM: Yeah. RC: And he has a high chair and he's near the wall. AM: Yeah. RC: And he's halfway out of the seat. Um, he's close, he's the closest one to the calendar. AM: Yeah. RC: And about that-much away from the um, a picture, and you can see just a little bit of the legs. That's all I can describe.	RC: The, the babies are sitting on a bed...and, and, the, it ain't the one where the little girl is sticking her feet out way apart. It's the other one.

* * * * *

Pretest	Posttest
CARTOON BOXES I	CARTOON BOXES I
RC: He's down on the trapeze and, um, his feet are up in the air, and, um, the um, the trapeze you can just see from the bottom. AN: Yeah. RC: And there's the um, there's green stuff down on the bottom. AN: Yeah. RC: And there's um, and Mutt has um, like the leg that you can't hop with, see, and it has like um, kind of like blocky things on it.	RC: It's the one before Jeff falls off and hits his head. His feet are like touching the black line... and the trampoline is kind of like up.
* * * * *	
CARTOON BOXES II	CARTOON BOXES II
RC: Bert's trying to push the bottom of the snowman up and Ernie's waving his hand. And, um, there's a tree right over Ernie's head and there's a building with like kind of a pointy top. AN: Yeah RC: And, um, there's a like l-l-little thing with grass and there's like a little thing - the tree's sticking out. AN: Yeah. RC: And um, Ernie has like you can hardly see his eyes. His eyes are together a lot.	RC: Bert's standing up on the biggest one...and putting little dots on it. BP: Oh, that's so easy. RC: And Ernie's handing him an apple.
* * * * *	
CARTOON BOXES III	CARTOON BOXES III
RC: The, um, guy's up in the tower and the girl's getting down.	RC: The man's pointing his head... alon- to, to the horse...and the girl's push- uh, taking in the, taking the braids out of her hair and pushing it down to the man...and he's disagreed... and he's putting his foot up on the castle like...and uh...it's the one before it's the one when he kicks, when he rides to the thing.

SCORING

Pretest	Posttest
Propositions rated + = 9	Propositions rated + = 17
Propositions rated 0 = 39	Propositions rated 0 = 23
Picture discrimination = 24	Picture discrimination = 31

Task: STL/Action-Sequence
Suject: #107
Status: High, Experimental

Pretest	Posttest
TEST DEMONSTRATION A	TEST DEMONSTRATION A
Put the red block in front of the yellow block and put the soldier on both blocks.	Drive the car behind the house and put the seal in the house.
Subject's response to A:	Subject's response to A:
JS: Put the red block in front of the yellow block, but leave a space that's as big as the block, and then put the soldier on it - on both of them.	JS: Drive the orange truck through... down, down this...um...road and through the green things, the green trees, and then... go up the other side, not the driveway...and then take the...um...then take the um... short cut and put it, park it right in back of the house, the car.

* * * * *

Pretest	Posttest
TEST DEMONSTRATION B	TEST DEMONSTRATION B
Take the elephant and hop him until he is between the trees and put the hippo behind the elephant.	Put the hippo on the car, and drive under the bridge and up to the red block, and put the hippo under the soldier.
Subject's response to B:	Subject's response to B:
JS: Take the hippo out from under the soldier and take the elephant and put him in the middle of the trees on the road and put the hippo in back of him.	JS: Take the hippo and put him on ...the...orange truck, on the truck, and..-drive it around and then under the bridge and park it where it used to be. And then you take the hippo out and put it...under the soldier, in that little opening.

* * * * *

Pretest	Posttest
TEST DEMONSTRATION C	TEST DEMONSTRATION C
Put the elephant and hippo in the car and drive it to the house. Put the elephant in the house and put the hippo under the soldier.	Put the elephant in front of one tree and the hippo in front of the other tree, and drive the car between them.
Subject's response to C:	Subject's response to C:
JS: Put the elephant and the hippo on the back of the car. Drive the car in front of the house. Put the elephant in the house and put the hippo under the soldier. The car facing the door.	JS: Put the elephant in front of the tree on the...um...left....And then put the hippo on the tree, in front of the tree on the right, and make them facing you. And then drive the truck down and park it right in the middle like...um...the road and then keep the...just drive it and park it at the end of the road of the driveway.

SCORING

Pretest	Posttest
Propositions rated + = 24	Propositions rated + = 24
Propositions rated 0 = 0	Propositions rated 0 = 2
Propositions rated - = 0	Propositions rated - = 0
Total FM score = 15	Total FM score = 14

Task: LL/Picture Description
Subject: #9
Status: High ,Experimental

Pretest	Posttest
PRODUCTION BOARD I	**PRODUCTION BOARD I**
ED: A tie and glasses, and they have white crosages, blue and red tie.	ED: They have...he has a mustache... and he has a red and white tie...and a white coat with black pants...and he...and he... GM: I got him. ED: And he, and the lady...has a green and blue and white...um...coat on and...and a green dress with a necklace.
PRODUCTION BOARD II	**PRODUCTION BOARD II**
ED: One has...a pink shirt on and a red skirt and she has short hair, and they're hugging. And the other one has a mustache and a white blouse and black pants.	ED: He has a black suit on...and a... a blue...and white tie...and he has a mustache...and... E: He's got it already. You were just going to say what the lady has. ED: The lady's ugly.
PICTURE BOXES I	**PICTURE BOXES I**
ED: She's looking at the dog and the dog is looking at her and she's laying down.	ED: She's looking towards you... and the...and the dog is looking towards her.
PICTURE BOXES II	**PICTURE BOXES II**
ED: He has dungarees on and his feet are up and he's... what do you call this? E: I don't know. You think of a name for it. ED: Bowling game.	ED: They're playing bowling... and...um...it's the boy's turn.

* * * * *

Pretest	Posttest
PICTURE BOXES III	PICTURE BOXES III
ED: The baby's looking towards you and one has a spoon in his mouth and the other one has a rabbit, and one, and the TV's on.	ED: The baby in the back seat, he's sort of crawling on the table, he's about to...um...the mother...and the the girl...is taking something out of her hand. E. The girl is taking something out of her hand? ED: Yep. The two boys are knocking heads.
* * * * *	
CARTOON BOXES I	CARTOON BOXES I
ED: He is going up in the air.	ED: Um...he's landing. His feet are up in the air and he's still on the mat...and, um...his hands are touching it.
* * * * *	
CARTOON BOXES II	CARTOON BOXES II
ED: He's putting on the mouth and one of them is standing on and the other one is giving him an apple.	ED: They have three...welll...they already have the three. And...Ernie's putting his hand up. And Bert is uh... pushing the three down.
* * * * *	
CARTOON BOXES III	CARTOON BOXES III
ED: She's putting her hair down and he's going to climb up it, and the horse is looking at him.	ED: He's holding her hair...the man's in the castle and..um...and he...and the man's holding her hair so she can get down. So, plus, he's getting down.

SCORING

Propositions rated + = 7	Propositions rated + = 14
Propositions rated 0 = 22	Propositions rated 0 = 16
Picture discrimination = 28	Picture discrimination = 31

Task: STL/Action-Sequence
Subject: #29
Status: Low, Control

Pretest

TEST DEMONSTRATION A

Drive the car behind the house and put the seal in the house.

Subject's response to A:

RG: Take your car, go around, go next to the house and drop the seal in the house.

Posttest

TEST DEMONSTRATION A

Put the red block in front of the yellow block and put the soldier on both blocks

Subject's response to A:

RG: Take the red block to the soldier ...out of the front of the door. Put the red block...well, take the soldier from the front of the door.

SB: I did that.

RG: Take the red block...put it where the wall...put it a little from the yellow block and put the soldier on the red block so he...he will be up in the air like.

* * * * *

Pretest

TEST DEMONSTRATION B

Put the hippo on the car, and drive under the bridge and up to the red block, and put the hippo under the soldier.

Subject's response to B:

RG: P-put your hippo on the car and go under the bridge and then put the hippo under the yellow blocks, take the red block, push it away and take the yellow block away and then push it and put the soldier on top.

Posttest

TEST DEMONSTRATION B

Take the elephant and hop him until he is between the trees and put the hippo behind the elephant.

Subject's response to B:

RG: Take your hippo- potamus and your elephant, and put it in the front of the car....first the hippopotamus then the elephant.

* * * * *

Pretest

TEST DEMONSTRATION C

Put the elephant in front of one tree and the hippo in front of the other tree and drive the car between them.

Posttest

TEST DEMONSTRATION C

Put the elephant and the hippo in the car and drive it to the house. Put the elephant in the house and put the hippo under the soldier.

Pretest	Posttest
Subject's response to C:	Subject's response to C:
RG: Put the car between the two trees. JB: Put the car between the two trees, how? RG: I mean, put the hippopotamus and the elephant between...beside the car...put the car on the road beside the two trees.	RG: Take the car...hold it...take the car, put the elephant and the hippo on it...and then drive it...to the house...drop the elephant in the house and take the hippo and put it under the thing where...the...soldier is.

SCORING

Propositions rated + = 14	Propositions rated + = 14
Propositions rated 0 = 4	Propositions rated 0 = 14
Propositions rated - = 7	Propositions rated - = 3
Total FM score = 10	Total FM score = 12

Task: LL/Picture Description
Subject: #56
Status: Low, Control

Pretest	Posttest
PRODUCTION BOARD I	PRODUCTION BOARD I
MF: She has a dress on. She has glasses.	MF: She has green on.

* * * * *

PRODUCTION BOARD II	PRODUCTION BOARD II
MF: She has a blouse on, a skirt and he has on a shirt, and tie and pants.	MF: He has a blue ti-, blue and white tie...and a black suit on.

* * * * *

PICTURE BOXES I	PICTURE BOXES I
MF: She's kneeling down, and she's petting the dog, and she's looking... left.	MF: On her right hand ...she has...two... a bracelet and a watch...and a ring.

* * * * *

Pretest	Posttest
PICTURE BOXES II	PICTURE BOXES II
MF: They're playing...they're laying down.	MF: The boy's touching a knob...and there's balls...and there's, um, pins facing them...they're laying down.

* * * * *

Pretest	Posttest
PICTURE BOXES III	PICTURE BOXES III
MF: They're on... He driving a train? They're all sitting down.	MF: They're in their highchairs... one's...getting on the other one's... highchair...one's standing up.

* * * * *

Pretest	Posttest
CARTOON BOXES I	CARTOON BOXES I
MF: He's jumping on the...he's jumping...he's going up and down.	MF: He's down on the trampoline.

* * * * *

Pretest	Posttest
CARTOON BOXES II	CARTOON BOXES II
MF: He's putting rai-...he's put them on top of each other, and, uh, Bert's putting raisins on his, uh, mouth.	MF: Bert's put- pushing the snowballs...down...and Ernie, Ernie isn't helping...Ernie's on the right side, and Bert's on the left side.

* * * * *

Pretest	Posttest
CARTOON BOXES III	CARTOON BOXES III
MF: She has lonf hair and she's in the castle, and the man's walking away. One of his foots on the castle, and his horse, horse is looking at him.	MF: Pri- Prince Charming is in the, um, window...and he's let Rapunzel down... and the prince is letting down her hair.

SCORING

Pretest	Posttest
Propositions rated + = 4	Propositions rated + = 6
Propositions rated 0 = 19	Propositions rated 0 = 14
Picture discrimination = 20	Picture discrimination = 25

APPENDIX II. A

A Core Curriculum
of
Communication Games

APPENDIX II. A.

CONTENTS

1. Bingo (3 games)
2. Go Fish (1 game)
3. Movie Director (4 games)
4. Who Am I? (2 games)
5. Object Arrangement (10 games)
6. Map Games (3 games)
7. Story Ordering (4 games)

A note on the materials cited in these games.

Most of the materials used in these games were original creations, developed or adapted for the specific purpose of a particular game. However, in preparing this report for public use we have attempted to identify commercially published material which we also used as an alternative, or which is comparable or similar to the materials actually used in the study. The sources of the materials cited are as follows:

For Bingo, Go Fish, and Who Am I? -- the picture cards from the Card Games series in Interaction: a Student-centered Language Arts and Reading Program, James Moffett, Senior Editor. Copyright 1973, Houghton-Mifflin Publishing Company, Boston, Massachusetts.

For the Map Games -- the cut-out paper figures from Press-Out Books like "Tiny Frontier Town", published by Whitman Books, Western Publishing Company, Inc., Racine, Wisconsin.

For the Story-Ordering Games -- cartoon strips from Mad Magazine Comics, Sesame Street Comics, and the Stories-without-words series from Little Golden Books, by Golden Press, New York. Published by Western Publishing Company, Racine, Wisconsin.

BINGO

General Structure of the Games:

The basic format of this communication exercise is similar to that of the popular game. There are a number of players split into two sides—on one side there is a single individual (the "caller") who directs the game, and on the other side there are the rest of the players who each play with their own array of cards. This version of the communication game requires that two duplicate decks of cards be used. The content of the cards is crucial because they must depict photos or pictures or drawings or designs that are not only different from one another in an absolute sense, but also contain a number of features that are similar across cards.

The different arrays of cards are arranged as follows:

The Players: Each player has in front of him 9 cards arranged in a 3 x 3 grid The identity of each item depicted on each card is unique, but each card also has something in common with a card in each of the other players' arrays.

For example, for 3 players the 3 arrays would be constructed as follows:

Player 'A"

1	2	3
4	5	6
7	8	9

Player 'B"

1	2	3
4	5	6
7	8	9

Player 'C"

1	2	3
4	5	6
7	8	9

Cards 1 'A', 1 'B', and 1 'C' might all be 3 different kinds of sea creature but all have in common the fact that they are all mammals.

The"Caller": The player who is the caller sits opposite the other players with a shuffled card deck that contains all of the cards in the players' arrays. For example, the first half-dozen cards he draws might look like this:

1A 5B 6A 3C 8B 5C 2B etc.

In each of the games described below, the caller draws one card at a time and describes it (without naming it) in such a way that one of the players can make a correct identification of that card in his own array. Each time a correct identification is made, the player turns that card face down. Whenever a player succeeds in turning 3 cards face down in a row, or a column, or diagonally, he gets "Bingo" and wins that round. For example, in the illustration above, if Player 'B' correctly identifies cards 5B, 8B, and 2B drawn by the caller, he will have Bingo.

There are two basic roles in these games. The Caller must describe the target cards and answer the players' questions. The Players must both decode and interpret the caller's messages and ask discriminating questions that will distinguish between two or three cards of similar content. No guessing, labelling or pointing is allowed prior to correct identification.

BINGO GAMES

Game I

Videotape Reference: VT #5, VT #6, and Demo. Tape.

Number of Players: Teacher and 3 children.

Age Range: All elementary grades.

Type of Language Elicited: Descriptions, discriminating questions, simile, metaphor, allusions. Also, depending on the content-area depicted by the cards (zoology, botany, wild life, etc.) frequent use is made of specialised vocabulary. Also the children are encouraged to formulate hypotheses and test them with intelligent questions.

Materials: Duplicate decks of cards with content and format pre-arranged, as described in the introduction above. For 3 players and 1 caller 2 sets of 27 cards each will be used.

Procedure and Aim of the Game: When the children are first introduced to this game it is best for the Teacher to play the "Caller" and the children to be the players. The children are instructed in the rules of the game, and the teacher should NOT begin the first round until she has made sure that the children understand that there is an overlap of content across the 3 arrays of the players. It should be emphasized that, because each child has cards that are similar (but not identical) to his neighbors', he must inspect not only his own array but also his partner's each time he hears the target card described by the "caller" (teacher). He must ask questions for further information when he is not sure since guessing is to be discouraged as much as possible. The children should be told to ask questions when they want to distinguish between two or more cards that seem to be potential candidates for the right answer (e.g., "does it have red spots on its back?").

The teacher draws the first card and the game begins. The first round ends when a player gets a Bingo. At the end of the first round the Teacher can (at his or her discretion) designate one of the children as the caller, while he becomes one of the players (or continue as caller himself until he is sure that the children understand the basic rules of the game). Whenever the Teacher becomes a player and/or a new round begins, the players all change positions so that they have a different array in front of them each time

As many rounds can be played as the children have patience and motivation for. A game of 4 rounds would give each child a chance to play both Encoder and Decoder once.

Role of the Teacher: The teacher's role is central here for modelling both adequate descriptions and question-asking strategies. As an Encoder (caller) she will have the opportunity to introduce rich vocabulary into her descriptions as well as to draw upon metaphor, simile and allusion to enrich and color the content of her message. As a Decoder (player) she will model question-asking techniques that focus on the similarities and differences between cards, and demonstrate the use of categorical questioning that can help discriminate and differentiate between competing cards.

Cognitive Discrepancy and Feedback: Feedback is immediate and continuous in this game, fo both caller and players. This game also has the added advantage of exposing each individual child to the question-answer dialogues of other players. This means that he will not only have the benefit of a variety of language models, but he will also gain practice in deduction and hypothesis formulation and testing since he will be able to take prior descriptions and questions into account in contructing his own subsequent questions.

"Go Fish Games"

Game I: "Feature Fish" with Dragon Pack

Number of players: Optimal number is 3-4 players.

Age Range: 6-10 yrs., for the kind of material used in this version of the game. If more complex material could be constructed then the game might still be appropriate for children over 10.

Type of Language Elicited: Question and answer dialogue. Descriptions, categorizations and classifications of attributes and features. Since the players can focus upon the absence as well as the presence of certain features, the logical language of conjunction, disjunction, inclusion and exclusion is also used.

Materials: Since the game is modelled after the children's game "go fish", it is played with a deck of cards which depict a variety of "suits" or representations. For this version of the game we use the "Dragon Pack" deck of cards from the Houghton-Mifflin Interaction series. Each of these cards depicts a cartoon drawing of a dragon. All of the dragons have certain similar characteristics e.g. size, shape, and color. However, each dragon is distinctive because it is made up of a combination of features, each of which varies along two dimensions-presence or absence. For example, a dragon can be with/without (±) tail, ± horns, ± breathing fire, ± tail, ± claws, etc.

It is these kinds of similarities and differences between dragons that allow the players to "match" or pair two dragons on the basis of certain features which they hold in common.

Procedure and Aim of the Game:
"Two-feature Fish": The aim of the game is the same as in the traditional "go fish" game, to make a match with two cards in your hand in order to win a trick or a point.

The procedure for this two-feature version of "go fish" follows three stages- first demonstration of the feature content of the cards, next modelling by the teacher of the rules of the game, and finally a complete round of playing the game.

1) The game is best begun with the teacher as model and 2 or 3 children as players. The teacher picks two cards from the deck to show to the children who are asked to indicate what the two dragons depicted thereon have in common. A third card is introduced and the teacher elicits more comparisons of the dragons from the children. As a general rule here, the teacher should try to play down certain obvious, non-critical features e.g., color, shading, wrinkles, etc., and, instead, have the children concentrate on the major features (horns, claws, tails, fire, etc.)

b) If the teacher decides there is no need to set a limit on the number of turns, she can allow players to win up to three successive turns before having to yield to the next partner.

Whether option A or B is chosen, a player's 'turn' always terminates after he has drawn a fresh card from the pack.

9) The game continues around all the players for as many rounds as are necessary to exhaust the center pack of cards and all possible matches have been made. As the game proceeds, some players will find themselves holding only one card, while others have two, when it becomes their turn to ask. In either case, each player must request a match from a partner, and, it is only after he hears a 'go fish' that he is allowed to draw from the center deck.

10) If a competitive element is desired, the winner of the game is the player who has made the most number of matches (or pairs) at the end of the game.

Complexity: The complexity of this game can be made appropriate to particular grade levels by manipulating two dimensions of the cards- *number of features*, and *feature presence vs. feature absence*-either separately or simultaneously.

1) **Number of Features**:- the game can be made easier or harder by reducing or increasing the number of features necessary to make a match. The above version, for example, can be increased to "three feature" or "four feature" fish by requiring the children to make matches of 3 or 4 characteristics of dragons at a time.

This added requirement can be introduced as a new session of the game, or it might even be introduced in the middle of a game to avoid monotony.

2) **Feature Presence vs. Feature Absence**: Once the children become familiar with the characteristics that the dragons overtly possess, they can be made aware of the binary dimension (present/absent) of the features by pointing out that dragons may sometimes *lack* certain features as well as possess them. Once the children come to understand this they can play a whole new round of the game, focussing on *absent* features, asking e.g., "Do you have a dragon without horns, and no tail?"

Finally, at the most complicated level of all, a version of the game can be played that combines the use of both present and absent features, e.g., "Do you have a dragon that breathes fire and has wings, but does not have a tail?" In this way the players learn how to construct messages that not only combine attributes but also impose qualifications and conditions that attest, in a more complete way, to the unique identity of the target dragon.

Cognitive Discrepancy and the Role of Feedback: In a small group, face-to-face game like this, feedback for the players is usually instantaneous. There is a constant match-up process going on between perceptual representations and verbal descriptions, whether it is with the Encoder who has to translate his chosen features into a verbal question form, or with the Decoder who has to match that question with the preceptual attributes of the cards in his hand.

MOVIE DIRECTOR

Game 1

Number of Players: Teacher and one child.

Age Range: All elementary grades.

Type of Language Elicited: Instructions, directions, use of terms for spatial orientation, bodily parts, part-whole descriptions.

Materials: A number of drawings or magazine pictures of people in different poses.

Procedure and Aim of the Game: The Teacher explains to the child that he is going to instruct the child how to act and stand just like a director in a movie. He shows the child examples of the different poses he is going to choose from. Then, without showing the child, the Teacher selects one card and begins to instruct the child how to stand just like the person in the pictur

Once they have completed the pose successfully, they exchange roles, with the child now instructing the teacher.

Cognitive Discrepancy and Feedback: One advantage of this game is that the Encoder receives immediate feedback about the effects of his messages on the Decoder's actions.

Modelling and the Role of the Teacher. Especially for younger children, the role of the teacher as a model is essential in this game. As an Encoder, the teacher can demonstrate the use of clear, accurate descriptions and how to give distinctive, non-ambiguous instructions. The teacher's role as a Decoder is probably even more important since he will be able to formulate questions intended to clarify the child's messages and require the Encoder to produce less ambiguous directions.

N.B. In order to assist the Encoder to verbalise some basic distinctions like "left/right" or "front/back" from his own perspective, it is often helpful to provide the child with some standard reference points to orient himself with, e.g., putting a watch or rubber band on the Decoder's left or right wrist.

Game 2

This game is identical to Game 1, except that it is played with two children, who take turns encoding and decoding the poses.

Game 3: "Live" posing.

This is idential to Game 1, with the following exception. The Teacher and 2 or 3 children between them think up some poses that can be modelled and photographed. Each child gets a turn to be the Encoder who instructs another child how to take the pose that he (the Encoder) has thought up. The Decoder holds his pose until a Polaroid Photo is taken of his position. This photo is then used in later games as a model for posing.

N.B. The children can get practice in constructing (and photographing) poses that involve one, two or three children.

WHO AM I?

Game I: "Categories," with Array

Number of Players: Can be played with as few as two or as many as five or six. However, in order not to have lengthy waits between "turns," it is probably best played with four children.

Age Range: Can be used throughout the elementary grades. The interest and appeal of the game is partly a function of the complexity of the materials, so the game can be made appropriate for any age level simply by choosing the right kind of materials.

Type of Language Elicited: In the "asking" version of the game, question-asking and use of classificatory language is required. In the "telling" version, the partners have to provide oblique descriptions and categorizations of groups and sets of object or pictures.

Materials: Two identical packs of cards or sets of pictures (e.g., the "On the Move" card decks about different forms of transportation, from the Houghton-Mifflin "Interaction" series). The content of the cards must represent a number of overlapping sets of categories or classes (e.g., vehicles that are airborne vs. vehicles that are landborne).

Procedure and Aim of the Game: One complete deck of cards is laid out in an array in front of the children. The array can be dealt randomly, but it is suggested that, for beginning levels of this game, the teacher group similar cards together in recognizable patterns (e.g., all the flying vehicles together, all the land vehicles together, etc.).

Each child in the team is then given one card from the duplicate deck to stick on his forehead so that each can see the other's but not his own. These "target" cards are pre-selected to represent different groups or categories in the array.

Asking Version: One child is designated as "it" or the target, and the game begins. Probably for beginning levels of the game it would be more instructive for the teachers to "*model*" the initial role of questionner to show the children the rules of the game and alert them to the existence and use of the categories in the materials.

The *aim* of the game is for the target child to ask each partner in turn "Who Am I?", until he can correctly identify his own picture. He does this by asking specific questions of each partner who can respond with only yes or no answers. The target child, through this kind of dialogue, has to gradually reduce the number of categories in the array that his picture belongs to until, by a process of deduction and elimination, successful identification is made.

Telling Version: Same as above except that the partners "tell" the target child who he is by providing clues and hints and oblique descriptions about the target picture. Labelling or naming of the pictures is not allowed.

Modelling by the teacher can also be done in the telling version of the game, with the teacher being one of the partners who is providing clues for the target child. This can be a very effective technique since the teacher can demonstrate effective communication to the other players through choice of vocabulary, grammatical structur of the message, sophisticated references and allusions, similes or metaphors, and so on.

Game II: Categories without array.

Versions: Two versions, either an asking version where the child questions his partners, or a telling version where the partners give clues to the target child.

Number of Players: 4 to 5.

Age Range: 6 years to teens.

Type of Language Elicited: Categorical questionning, analogical descriptions, logical statements of inclusion and exclusion. There will also be an educational, instructional content to the dialogues that will depend on the kinds of representational material used in the games (e.g., zoology, botany, technology, etc.

Materials: No array is used in this game, but each child has on his forehead a picture that is related in some systematic way to each of the other players' pictures. For example, the game could use the "Life in the Sea" cards or the "North American Animals" cards from the Houghton Mifflin "Interaction" series. The complexity of the game will depend on the selection of target pictures. These may all have one thing in common (e.g., all fish), or they may represent two distinct classes of animals (e.g., dogs vs. cats). The game may be further complicated by adding further dimensions and attributes. For example, a matrix classification can be achieved by choosing pictures which contrast canine vs. feline and wild vs. domestic. In this way, target pictures can be carefully selected and matched in order to elicit dialogue about different levels and hierarchies of classification and interrelationships.

Procedure and Aim of the game:
--Asking: The children are instructed that the pictures are related in some way and that non-target pictures can provide clues to target pictures. The target child goes to each partner in turn asking questions about his own identity. By looking at the others' pictures, he is enabled to deduce, to a certain extent, what sort of an animal, fish, etc. that he is. The materials, therefore, provide a starting point and a focus for his questions which, if they are to be successful, must become increasingly specific.

--Telling: As above, except that the partners, in turn, each "tell" the target child who he is by providing clues and hints.

Modelling: To introduce this game, the teacher can participate in the early rounds, to model card, demonstrate the kinds of questions or descriptions that are most effective for solving the communication problem.

This game should be played for two or three rounds with different sets of pictures in order to maximize the cumulative learning effects that take place, as each child benefits from the discussion and style of questionning that has gone on in previous rounds.

The Role of the Teacher: This is essentially the same ~~as in~~ the other versions of Who Am I?. The teacher helps the children keep track of the game, and keep their questions focussed and not wandering too far afield. She draws their attention to categories or classifications that may have gone unnoticed, and generally keeps the questionning or clue-giving, moving towards the goal of correct identification.

OBJECT ARRANGEMENT

Series I (Toys)

Game I. Free construction prior to verbalization in "screen" situation.

Number of Players: Two- teacher and child

Age Range: For lower elementary grades, and for older children who might need this level of exercise as a practice before proceeding to more difficult material.

Type of Language Elicited: Descriptions, directions, instructions. Use of spatial and relational terms and prepositions.

Materials: Variety of simple toys containing people, animals, buildings, vehicles, etc., which can be used to create simple scenes and arrangements. (e.g. house, tree, tower, horse, car, truck, soldier). Duplicates of all toys are needed for both partners.

Procedure and Aim of the Game: The child plays the role of Encoder here. The child and teacher are seated on either side of a table, facing each other but separated by a low screen such that they can see each others faces but not the table top. She is given 3 or 4 objects from the toy assortment (e.g. house, horse, car, truck), and is instructed to arrange them any way she wants to on the table in front of her, to make a little scene.

The aim of the game is for the child to describe the arrangement he has built to the teacher so that the latter can build the identical arrangement with a similar set of toys on his side of the table.

Role of the Teacher: Since the teacher is in a genuine role of ignorance about the child's arrangement, he should attempt to improve the quality of the child's message by asking pointed questions that will force the child to rephrase and restate inaccurate, ambiguous or imprecise messages. By so doing, he will help the child to hone his messages to the required degree of precision and specificity.

Cognitive Discrepancy and Correction of Messages: After the child has completed his descriptions, he is brought around to stand behind the teacher to see if the latter's arrangement conforms to the child's. If a discrepancy exists between the child's and the teacher's arrangements, the child remains behind the teacher and tells him "over the shoulder" what changes need to be made.

Materials: Same toys as in Game 2, but this time the photos are of different arrangements or groupings of the objects.

Procedure: The children take turns playing the encoder, each using a different photo for instructing his partner.

Series I

Game 4: Free construction prior to verbalization in "screen" situation.

This game is identical to Game 1 with the following exceptions:

Number of Players: Two children, preferably those partners who have previously played Game 3.

Materials: Same as Game 1

Procedure: In this version all of the toys are used. One child is designated as the first encoder, and is instructed to build a little village with the toys in any arrangement he likes. After arranging the toys the encoder then describes the assembly to his partner so the latter can duplicate it on his side of the screen.

After the description is completed (and any corrections made "over the shoulder") the children exchange roles and the new encoder builds and describes a fresh arrangement.

14

OBJECT ARRANGEMENT

Series II (familiar objects)

Game 5 : Construction of arrangement, with help of photo, during verbalization, in "screen" situation.

Number of Players: Two children who each take turns playing Encoder.

Age Range: All elementary grades, provided material is made appropriate to the age level.

Type of Language Elicited: Descriptions, directions. instructions, with us of spatial, relational terms and prepositions (N.B. In this particular version, the requirement that the Encoder verbalise *during* construction means that he has to "think out loud.")

Materials: For this simpler version of the game, duplicate sets of toy objects and non-toy familiar objects can be used, e.g. doll, toy cup, toy chair fork, pencil, paper cup etc. In addition, Polaroid photos of some kind of "still life" arrangements of the objects. For this version, the photo scene should include not only neutral arrangements (e.g., pencil inside paper cup) but also arrangements that involve some intrinsic or meaningful relationships (e.g., 'the doll is sitting on the chair drinking from a cup').

Procedure and Aim of the Game: The two children are seated on either side of a screen. Both are given the same set of objects. The chil who is Encoder is also given a photo of an arrangement to build He is instructed that *while* he is building the model in the photo he is to instruct his partner to build along with him so that they will both end up with identical arrangements. After the first game is completed, the Encoder is given a fresh photo of another arrangement to describe

Cognitive Discrepancy and Correction of Messages: In each game after the Encoder has completed his message, he is told to come around an stand behind his partner, holding the photo to inspect the Decoder's arrangement. If there is any discrepancy between the standard arrangement and the partner's assembly, the Encoder is to give the additional verbal instructions necessary to correct the situation.

Series II

Game 6B: Constructing an arrangement (with photo and materials) while verbalizing, in a "screen" situation.

Number of Players: Two children, one of whom has just played Game 6A

Age Range: Same as Game 6A.

Type of Language: Same as Game 6A.

Materials: Duplicate sets of the same objects and the same photo that was used in Game 6A.

Procedure and Aim of the Game: The two children are seated at a table on either side of a screen. Each has scattered in front of him the set of objects. The child who has played Game 6A prior to this is designated as Encoder. He is given the photo and instructed to perform two tasks simultaneously- he is to arrange his materials to match the model in the photo and, at the same time, tell his partner what he is doing so the partner can correctly duplicate the arrangement on his side

Cognitive Discrepancy and Correction of Messages: When the assemblies have been completed on both sides the Encoder comes around with his photo to stand behind the Decoder and give any additional instructions necessary to correct any inaccuracies in the Decoder's arrangement.

17

OBJECT ARRANGEMENT

Series III (Blocks, abstract shapes)

Game 7: Pre-construction of blocks before verbalization, in "screen" situation.

Number of Players: Two children who both take turns playing Encoder and Decoder.

Age Range: All elementary grades

Type of Language: Descriptions, directions, instructions. Use of spatial and relational terms, and prepositions

Materials: Two sets of half a dozen wooden rectangular blocks

Precedure and Aim of the Game: This is a "free form" construction task The two children sit on either side of a screen with the blocks in front of them The Encoder is instructed to build anything he wants with the blocks. After he is finished he is told to describe his arrangement to his partner in such a way that the Decode can replicate the assembly on his side. At the end the children change roles, and the game is repeated

Cognitive Discrepancy and Correction of Messages: When the Decoder has finished building the model, the Encoder comes around to see if the Decoder's assembly corresponds to his own. If there are any discrepancies, the Encoder returns to his seat and gives the necessary additional instructions across the screen.

Series III

Game 9: "Over the Shoulder" Instruction, using one photo or serial photos.

Number of Players: Two children.

Age Range: All elementary grades.

Type of Language: Same as Game 8.

Materials: A number of wooden blocks or geometric shapes, a photo of a standard arrangement of these blocks, and a set of serial photos showing the successive stages of construction up to and including the final model.

Procedure and Aim of the Game: The child who is Decoder sits at a desk with the blocks scattered in front of him. The Encoder stands behind the Decoder and instructs him "over the shoulder" what arrangement to build, using the photo (or photos) as a model. N.B. This version of the game can be played in an easier or harder way depending on the Encoder's experience and abilities. In the easier version, the Encoder is giventhe set of serial photos and, usi one photo at a time, talks the Decoder through each of the successive building stages up to the final stage. In the harder version, the Encoder is given only the single photo of the completed assembly, and it is left up to him what strategy to devise for ordering and sequencing the steps of the construction to be carried out by the Decoder.

The Teacher should prevent the Encoder from trying to communicate via pointing or gesturing.

Game 10: Description from Photo in "Screen" situation.

Number of Players: Two children who take turns playing Encoder and Decoder.

Age Range: All elementary grades, but younger children will probably need to have had prior training with the games in Series I and II.

Type of Language: Same as in Game 8

Materials: Duplicate sets of wooden blocks (e.g., cylinders or dowels of different lengths). Photos of standard arrangements of the blocks in some sort of abstract array.

Procedure and Aim of the Game: The children are seated on either side of a screen, and only the Decoder has the blocks scattered in front of him. The Encoder is given a photo and told to instruct the Decoder how to build the model shown in the photo.

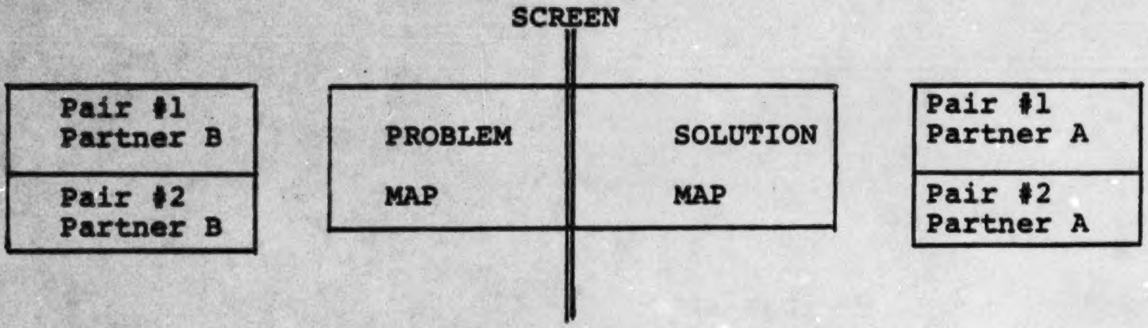
The children take turns at giving the descriptions, with a new photo being used each time.

Cognitive Discrepancy and Correction of Messages: At the end of each game, the Encoder is brought around to stand behind the Decoder, holding the photo, to check the correctness of the Decoder's arrangement against the photo model and give additional instructions when necessary.

Map Games

Rationale

The general idea in these map games is to take a problem situation - getting from point A to point B - and engage children as partners in a communication exercise aimed at achieving a cooperative solution to the problem. The problem situation is created by constructing a map that contains not only starting points, routes, landscapes and goals, but also problem points that need to be overcome and which require specific solutions. The complexity of the game (and the need for verbal dialogue) is created by splitting the basic map into two versions - one that contains the problems, and one that contains the solutions. The usual way to play the game is in a "screened" situation where a version of the map is placed on either side of the screen. Four players are chosen and divided into two teams of pairs. Having divided the map in two, the pairs are then also divided with a partner from each pair assigned to different sides of the screen thus:



Partners A & B on the same team must communicate with each other about the different kinds of information they each encounter on their respective maps. A competitive element is introduced into the game as each pair tries to reach "home" before the other pair.

Generally, therefore, neither side possesses all the information necessary to make progress along the routes, to avoid danger and reach the goal safely. Each child therefore is required to use his language skills for exchanging descriptions and instructions with his partner, and a question - and - answer dialogue must be maintained if a satisfactory and cooperative solution is to be reached.

The kind of language elicited by this type of conditional or contingent communication game can be quite rich depending on the complexity of the version played. The language can range from simple descriptions all the way to conditional propositions of the "if/then" variety depending on the particular contingencies encountered in the particular game.

/3

Frontier Map Game

Game II - "Problems and Solutions"

Number of Players: Can be played with as few as two or as many as six. However, it is probably best played with four children.

Age Range: 7-14. The interest and appeal of the game is partly a function of the complexity of the boards, so the game can be made appropriate for any age level simply by choosing the right kind of materials.

Type of Language Elicited: Descriptions and instructions, together with explanations and judgements based on the kinds of problems and contingency situations encountered along the route.

Materials and Format of the Game: Dice, plastic cowboy playing pieces, and two map boards (A & B) which come in two versions: In the simple Version I, all of the different paths are differentiated by shape and color (e.g. paths constructed of orange circles vs. paths made of blue squares); in the harder Version II the paths are all of the same shape and color.

Regardless of which version is used Boards A & B both have certain standard features. Both Boards are identical in that they both have the same paths and similar land scape features. They differ, however, in that Board B (the "problems" board) has a number of buildings (saloon, hotel, barbershop, etc.) at strategic locations, together with a number of "problem" spots along the route which require "solutions" - e.g. encountering a tin-star on the path necessitates a detour to the sheriff's office.

Also included with the materials is a legend explaining the problems and their solutions.

Procedure and Aim of the Game: The pairs of players take turns directing their partners to destinations on the board using discriptions of the route and landscape. 'Teams' and 'partners' line up on either side of the screen in the manner indicated in the diagram in the introduction. Thus the object of the game for Partner 1A is to get himself and his Partner 1B from the "start" to the "bank" to deposit some gold faster than Partners 2A and 2B.

Child 1A and child 2A each take turns throwing the dice, choosing a route and instructing their respective partners, as follows:

1) To begin, the dice is thrown by each player at Board A, and the higher scorer goes first.
2) The first player at Board A throws the dice and moves his playing piece the appropriate number of spaces along a route of his choosing. During or after these moves he communicates what he is doing to his partner on the other side of the screen so the latter can make the identical moves on Board B.

In version 1 of the Boards a typical message from A to B might be "move three spaces on the path with the orange circles." In the

15

Map Games

Game III - Problems and Solutions, with RISK Cards

This game is identical to Game II in all respects with the following exceptions.

Materials: In addition to the regular materials, a set of "RISK Cards" are included. These cards are used (face-down) to cover the "?" (query) squares on Board A. The face side of each of these RISK Cards depicts a situation which, depending on the Card and its location on the Board, may have positive or negative consequences for the player who chooses it. For example, a negative RISK Card may denote a person hungry for an apple, with the instruction that the player must retrace his steps to the Apple tree. On the other hand, a positive RISK card may allow the recipient to take extra moves.

Procedure: This game is played in two versions: The first introduction to this game should be with Version 1 where the use of the RISK Cards is compulsary. After that can be played Version 2 where the cards are optional.

The players on Board A get to use the cards to provide additional route instructions for their partners. In Version 1 players must continue on the roads they have been taking when those roads end at or traverse a RISK Card. They are not allowed to detour the card. Landing on, or crossing over a RISK Card requires that that card be picked up and read, and its consequences followed, whether positive or negative. In Version 2, players are allowed either to bypass and detour the RISK Cards or to play them and take the RISK involved.

To prevent the players from learning the content of the RISK Cards from their position, the cards should be shuffled and randomly placed prior to each game.

STORY ODERING GAMES

Game I Cartoon Boxes - "Race Cars"

Number of Players: Two children

Age Range: Lower Elementary grades

Type of Language: Story-telling, sequential narrative and discourse

Materials: Duplicates of a Mad Magazine cartoon strip about "Race Cars." One strip is arranged in it's natural sequence on a strip of cardboard. The other strip is cut into 3 separate "boxes." The game is played with a Screen.

Procedure and Aim of the Game: The aim of the game is for the Encoder to tell the correct story sequence to a Decoder who has to rearrange his cartoon boxes into the proper order.

Both children are seated on either side of the screen. The teacher gives the cardboard cartoon strip to the Encoder first, and gives nothing to the Decoder at this point. The teacher explains to the Encoder that he has to tell the story to the Decoder who has all the parts to the story but in a mixed up order. The Encoder, therefore, has to tell the story well enough to enable the Decoder to arrange his cartoon boxes correctly.

The teacher gives the Encoder enough time to study the story and only when the Encoder indicates that he is ready does the teacher give the separated cartoon boxes to the Decoder in a mixed up order.

The game ends when the Decoder correctly sequences his cards.

Game II Cartoon Boxes - "Giving the Dog a Bath"

This game is similar to Game I with the following exceptions.

Materials: The Cartoon Boxes come from a Sesame St magazine strip about bathing a dog. Duplicates of the strip are cut into 5 separate cartoon "boxes", and each player receives an identical set of boxes. The game is played with a Screen.

Procedure: The game is introduced by the teacher as in Game I, except that, this time, the Encoder is given his loose assortment of cards in a mixed up order, and not pre-arranged.

Therefore his first task is to sort out the story sequence before he begins to describe it to the Decoder. Once again, the Decoder does not receive his mixed-up set of 5 cards until the Encoder is ready to begin. At that point the teacher lays out the 5 cards in front of the Decoder in a random order.

Game IV: Round Robin Story-telling (unnumbered sequence) "Camping out in the Playhouse"

This game is similar to Game III with the following exception.

Materials and Procedure: As before, there are 8 picture cards but this time they are not numbered, so the children will have to cooperate and compare cards in order to correctly sequence the pictures in the story. the teacher distributes the cards randomly to each child, taking the first card for herself in order to begin the story. This is a harder version of Round Robin because it requires that the children make a cooperative interpretation of each step of the story sequence. The overall story, therefore, will truly be the result of combined effort.